Université de Montréal

Identification and characterization of new biomarkers in aggressive subtypes of breast cancer

par

Einas Yousef

Programmes de biologie moléculaire

Faculté de médecine

Thèse présentée à la Faculté de médecine

en vue de l'obtention du grade de Ph.D
en biologie moléculaire
option biologie des systèmes

Mai, 2015

© Einas Yousef, 2015

Résumé

En 2015, la récidive tumorale et les métastases du cancer du sein demeurent une cause importante de décès à travers le monde. Toutefois, ces cancers sont souvent hétérogènes car en dépit d'un phénotype similaire, l'évolution clinique et la réponse au traitement peuvent varier considérablement. Il y a donc un intérêt évident à identifier et à caractériser de nouveaux biomarqueurs pour permettre classer les tumeurs mammaires dans des sous-groupes plus homogènes. Notre hypothèse est que chaque cancer mammaire possède des caractéristiques distinctes au plan des altérations du génome et des profils d'expression géniques et que ces changements se traduisent cliniquement par une prédisposition à former des métastases ou à répondre ou non à la chimiothérapie et aux thérapies ciblées. Dans le cadre de nos travaux, nous nous sommes intéressés aux sous-types agressifs de tumeurs mammaires et notamment les cancers de type triple négatif. Nous avons aussi tenté d'identifier des marqueurs capables de distinguer l'une de l'autre les tumeurs de type luminal A et luminal B.

Pour ce faire, nous avons d'abord utilisé une stratégie in silico à partir de données publiques (micro-puces d'ADN et séquençage de l'ARN). Nous avons ensuite construit sept micro-matrices tissulaires (TMA) provenant de tissus mammaires normaux et tumoraux fixés à la formaline et enrobés en paraffine. Ces outils nous ont permis d'évaluer par immunohistochimie les niveaux d'expression différentielle des marqueurs suivants : ANXA1, MMP-9, DP103 et MCM2. Ceux-ci ont été comparés aux marqueurs usuels du cancer du sein (ER, PR, HER2, CK5/6 et FOXA1) et corrélés aux données cliniques (survie globale et métastase).

Nos résultats indiquent que ces nouveaux marqueurs jouent un rôle important dans l'évolution clinique défavorable des tumeurs de haut grade. Dans un premier article nous avons montré que l'expression d'ANXA1 est dérégulée dans les cancers de type triple-négatif et aussi, dans une certaine mesure, dans les tumeurs HER2+. Nous croyons qu'ANXA1 permet de mieux comprendre le processus d'hétérogénéité tumorale et facilite l'identification des tumeurs de haut grade. Nous proposons également qu'

d'ANXA1 stimule la transition épithélio-mésenchymateuse (EMT) et la formation des métastases.

Dans un second temps, nous avons montré que les niveaux d'expression de MMP-9 reflètent la différenciation cellulaire et corrèlent avec les sous-types de cancers mammaires ayant un mauvais pronostic. Nous estimons que MMP-9 permet de mieux comprendre et d'identifier les tumeurs mammaires à haut risque. De fait, la surexpression de MMP-9 est associée à une augmentation des métastases, une récidive précoce et une diminution de la survie globale.

Dans le cadre d'un troisième article, nous avons montré que la surexpression du marqueur de prolifération MCM2 s'observe dans les cancers triple-négatifs, HER2+ et Luminal B par comparaison aux cancers luminal A ($p< 0.0001$). Nos résultats suggèrent qu'en utilisant un seuil de 40% de noyaux marqués, nous pourrions distinguer l'une de l'autre les tumeurs de type luminal A et luminal B. Cela dit, avant de pouvoir envisager l'utilisation de ce marqueur en clinique, une étude de validation sur une nouvelle cohorte de patientes s'impose.

En somme, les résultats de nos travaux suggèrent qu'ANXA1, MMP-9 et MCM2 sont des marqueurs intéressants pour mieux comprendre les mécanismes physiopathologiques impliqués dans la progression tumorale et le développement des métastases. À terme, ces nouveaux marqueurs pourraient être utilisés seuls ou en combinaison avec d'autres gènes candidats pour permettre le développement de trousses « multigènes » ou d'essais protéomiques multiplex pour prédire l'évolution clinique des cancers mammaires.

Mots-clés: Cancer du sein humain, biomarqueur, l'analyse silico, matrices tissulaires, ANXA1, MMP-9, MCM2, Ki-67

Summary

In 2015, breast cancer remains a leading cause of death among women worldwide due to relapse and metastases. However, mammary tumors are known to be heterogeneous in terms of their clinical course and response to treatment, despite a seemingly similar phenotype. There is therefore an obvious need to identify and characterize new biomarkers of progression in breast cancers so that each tumor can be properly classified. Our hypothesis is that each breast cancer has its own set of genomic abnormalities or altered pattern of gene expression that can explain the aggressiveness of each tumor, its ability to metastasize and its response to chemotherapeutic agents or other forms of targeted therapies. In this study, our aim is to identify and characterize new biomarkers with prognostic value in aggressive subsets of breast cancer focusing primarily on triple-negative tumors and luminal B breast cancer.

To achieve those aims, we conducted an *in silico* search from public databases of DNA microchip and RNA sequencing data. We next constructed seven tissue microarrays (TMA) using paraffin blocks from human breast cancer along with normal breast to examine the differential expression of new putative markers: ANXA1, MMP-9, DP103 and MCM2. Expression levels measured by immunohistochemistry were then compared to other conventional markers of breast cancer (ER, PR, HER2, Ki-67, CK 5/6, FOXA1) and correlated with clinical data (overall survival and metastasis).

By comparing the relative expression of these markers in human breast tumors we were able to pinpoint the important role of ANXA1, MMP-9, DP103, and MCM2 in aggressive tumor subtypes recognized for their poor clinical course. Firstly, we have shown that ANXA1 expression is severely deregulated in high-grade breast cancers including triple-negative and, to some extent, HER2-positive breast cancers. In addition, our results also indicated a possible role of ANXA1 in regulating EMT and breast cancer cell metastasis.

Secondly, expression of MMP-9 was found to mirror the degree of tumor differentiation and to correlate with breast cancers of unfavorable outcome. This implies that MMP-9 can help better characterize the biology of breast carcinoma and to identify

subgroups of high-risk breast tumors. In fact, we found that high levels of MMP-9 in tumors were associated with increased metastatic dissemination, early relapse and reduced survival.

Thirdly, we demonstrated that MCM2 is overexpressed in triple-negative, HER2 positive and luminal B breast cancer in comparison to luminal A breast cancer (*p-value* < 0.0001). Our findings support the notion that MCM2 can be used to distinguish luminal A from luminal B breast cancer based on a 40% index cut-point. However, an independent validation cohort is needed to confirm the clinical utility of MCM2.

Lastly, our results suggest that ANXA1, MMP-9 and MCM2 are valuable genes/proteins candidate that can help better understand the mechanisms involved in tumor progression and metastasis. One may also envisage their use, alone or in combination with other genes, in the development of a multi-gene panel or multiplex proteomic assay to predict clinical outcome and guide therapeutic decisions.

Keywords: Human breast cancer, biomarker, *in silico* analysis, tissue microarray, ANXA1, MMP-9, MCM2, Ki-67

Table of content

Résumé ..... **iii**

Summary ..... **v**

Table of content ..... **vii**

List of figures ..... **ix**

List of tables ..... **xii**

List of abbreviations ..... **xv**

Remerciements / Acknowledgment ..... **xxi**

CHAPTER I **1**

Introduction **1**

1. Hallmarks of cancer **1**

2. Normal breast tissue **5**

3. Breast cancer **7**

3.1 Incidence and heterogeneity of breast cancer 7

3.2 Histological classification of breast cancer 7

3.3 Molecular classification of breast cancer 9

3.4 Immunohistochemical markers as surrogate for molecular subtyping and its pitfalls 18

4. Cancer biomarkers **19**

4.1 Biomarkers: from discovery to clinics 20

4.2 Classifications of tumor biomarkers: 21

4.3 Problems associated with biomarkers 26

5. Biomarkers in breast cancer **29**

5.1 Established biomarkers in breast cancer 29

5.2 Multigene expression assays available in clinical setting 39

5.3 Proliferative biomarkers in breast cancer 45

5.4 Other emerging biomarkers in breast cancer 50

6. Objectives of this work **54**

CHAPTER II **55**

Deregulated Expression of ANXA1 in Human High-Grade Breast Cancers **55**

Abstract 57

Introduction 58

Material and methods 60

Results 67

Discussion 80

CHAPTER III **85**

MMP-9 expression varies according to molecular subtypes of breast cancer **85**

Abstract 87

Introduction 89
Material and Methods 90
Results 94
Discussion 108
Conclusion 111
Additional materials 112

CHAPTER IV 113

MCM2: an alternative to Ki-67 for measuring breast cancer cell proliferation 113

Abstract 115
Introduction 117
Material and methods 119
Results 124
Discussion 142
Conclusion 145
Additional materials 146

CHAPTER V 151

Discussion 151
Conclusion 159
Perspectives 159

APPENDIX 161

Appendix I: Panel of genes in MammaPrint assay (70-gene assay) 161
Appendix II: Panel of genes in PAM50 Breast Cancer Intrinsic Classifier™ assay (50-gene assay) 164
Appendix III: DEAD-box helicase DP103 defines metastatic potential of human breast cancers 166

Abstract 169
Introduction 170
Results 172
Discussion 201
Materials and methods 205
Supplemental information 209

Appendix IV: Identification of genes that could distinguish between luminal A and luminal B molecular subtypes of breast cancer using *in silico* analysis 246

References 261

List of figures

CHAPTER I:

Figure 1.1: Hallmarks and enabling characteristics of cancer…..…….....3
Figure 1.2: Structure of normal breast…….……..……..…6
Figure 1.3: Different histological subtypes of breast cancer…….……..…8
Figure 1.4: Different histological grades of breast cancer……...…….……10
Figure 1.5: Structure and different mechanisms of action of ER……..31
Figure 1.6: Structure and mechanism of action of PR……..……34
Figure 1.7: IHC scoring of HER2 expression……...……38
Figure 1.8: Role of MCMs complex in precise DNA replication…….49

CHAPTER II:

Figure 2.1: *In silico* analyses of ANXA1 expression in PAM50 molecular subtypes of breast cancer…..…….....…...69
Figure 2.2: Expression of ANXA1 in different types of breast cancer cell lines.….....71
Figure 2.3: Expression of ANXA1 in normal breast tissue and *in situ* carcinoma using immunohistochemistry and direct immunofluorescence……..…….72
Figure 2.4: Expression of ANXA1 in different molecular subtypes of breast cancer…...74
Figure 2.5: ANXA1 expression in breast cancer of different histological grades….……75
Figure 2.6: Genes that are positively or negatively correlated with ANXA1 in 756 breast cancer patients based on RNA-sequencing data derived from TCGA dataset……..…77
Figure 2.7: Correlation between ANXA1 and VIM, ESR1 and FOXA1 using RNA sequencing data…….……...…..…....78
Figure 2.8: Correlation between ANXA1, ESR1 and FOXA1 in different molecular subtypes of breast cancer using immunohistochemistry and tissue microarray……79

CHAPTER III:

Figure 3.1: *In silico* analysis of *MMP-9* mRNA expression in breast cancer subtypes….95
Figure 3.2: Expression of *MMP-9* mRNA in human breast cancer cell lines…..……...96
Figure 3.3: Validation of MMP-9 antibody specificity for IHC studies…..……...98

Figure 3.4: Expression of MMP-9 in normal breast tissue…..……...99
Figure 3.5: Overexpression of MMP-9 is associated with triple-negative, HER2-positive breast tumors and nodal metastases…….101
Figure 3.6: Overexpression of MMP-9 is associated with a higher incidence of metastases……105
Figure 3.7: Overexpression of MMP-9 is associated with shorter time to relapse and shorter survival after relapse. …….…….107

CHAPTER IV

Figure 4.1: *In silico* analysis of *MCM2* and *MKI67* mRNA expression in PAM50 molecular subtypes of breast cancer ……...……126
Figure 4.2: Patterns of MCM2 and Ki-67 expression in normal breast tissue……129
Figure 4.3: Overexpression of MCM2 is associated with high histological grade of breast cancer. ……....131
Figure 4.4: Intratumoral heterogeneity for Ki-67 labeling but not for MCM2, in breast cancer tissues. …….……135
Figure 4.5: Overexpression of MCM2 is associated with triple-negative, luminal B and HER2-positive breast cancer……137
Figure 4.6: A bimodal frequency distribution of MCM2 scores is detected in luminal breast cancer…….……...138
Figure 4.7: High levels of MCM2 are associated with a shorter latency to relapse…...140

Additional file 1: *In silico* analysis of MCM2, MKI67, ESR1 and PGR mRNA expression in PAM50 molecular subtypes of 597 breast cancers obtained from UCSC Cancer Genomics Browser……...146
Additional file 2: Validation of MCM2 antibody specificity for IHC studies…….147

Appendix III:

Figure 1: DP103 levels correlate with invasiveness and malignancy……..175
Figure 2: DP103 regulates invasive abilities of cancer cells……...180
Figure 3: DP103 regulates MMP9 expression and function……184

Figure 4: DP103 is an essential regulator of NF-κB signaling……190
Figure 5: Expression of active p65 (phospho-p65, Ser276) has a prognostic value……192
Figure 6: DP103 regulates NF-κB in response to multiple stimuli……..196
Figure 7: DP103 is a positive cofactor of TAK1-mediated IKK2 activation in MDA-MB-231 cells……...199
Figure 8: Schematic model based on our study showing a role for the RNA helicase DP103 in its ability to bind and stabilize TAK1 and thus activate NF-κB signaling in cancers. ……203

Figure S1. : Expression of DP103 in different molecular subtypes of breast cancer…..218
Figure S2: Expression of DP103 in different histological grades of breast cancer and its positive correlation with metastasis-associated genes…….220
Figure S3: Suprresion of DP103 decreases migratory ability of breast cancer cells…...222
Figure S4: Suppression of DP103 decreases invasive ability of breast cancer cells while its ectopic expression enhances cell invasion……..224
Figure S5: MMP9 predict survival and strongly correlate with DP103 levels in breast cancer. Also, suppression of DP103 expression decreases MMP9 gene expression…...226
Figure S6: MMP9 mediates effects of DP103 in invasiveness……228
Figure S7: DP103 regulates NFκB-dependent gene expression……..229
Figure S8: The helicase dead mutant GNT of DP103 retained its ability to induce invasion in MDA-MB-231 cells……..231
Figure S9: DP103 enhances TAK1-mediated IKK2 phosphorylation and hence NFκB activation……..233

Appendix IV

Figure 1: steps of selection of target genes that can distinguish between luminal A and luminal B breast caner…….247
Figure 2: Proliferation cluster according to MiSTIC that include the genes of interest…….248

List of tables

CHAPTER I:

Table 1.1: Parameters of histological grading of invasive breast carcinoma……..…...10
Table 1.2: Characteristics of different molecular subtypes of breast cancer…….……12
Table 1.3: Multigene expression assays available in clinical settings….……..41
Table 1.4: Panel of genes in Oncotype Dx™ assay (21-gene assay)…….……42
Table 1.5: Oncotype Dx™ recurrence score…..…….42

CHAPTER II:

Table 2.1: Clinico-pathological data of tissues used in TMAs.…….……63
Table 2.2: Antibodies used for immunohistochemistry…….……..……..65
Table 2.3: Immunohistochemical scoring system..……..……..66
Table 2.4: Expression of ER, PR, HER2 and ANXA1 in different types of breast cancer cell lines.……..…….…….70

CHAPTER III:

Table 3.1: Univariate analysis of different factors that could affect level of MMP-9 expression. ……...……...103
Table 3.2: Multivariate analysis model of MMP-9 that include metastasis, histological subtypes and molecular subtypes.……105

Additional file 1: Number of patients with or without metastasis associated with either high or low MMP-9 expression. ……112

CHAPTER IV

Table 4.1: Clinico-pathological data of tissues used in TMAs…….……..121
Table 4.2: MCM2 visual and automated scores in different molecular subtypes of breast cancer……..136

Table 4.3: Univariate Cox regression analysis of different factors that could affect OS, SAR and relapse.…….141
Table 4.4: Multivariate Cox-regression analysis of different factors that could affect OS and SAR.……..141

Additional file 3: Immunohistochemical scoring systems……...148
Additional file 4: Classification of molecular subtypes of breast cancer based on IHC surrogate markers. ……..…….150

Appendix I: Panel of genes in MammaPrint assay (70-gene assay)……...161

Appendix II: Panel of genes in PAM50 Breast Cancer Intrinsic Classifier™ assay (50-gene assay)……..…….164

Appendix III:
Table 1: Metastatic qPCR array……..……..……...183

Table S1: A) & B) Clinicopathological features (Singapore)…..……..……..235
Table S1: C) Clinicopathological features (Canada).……..……236
Table S1: D) Clinicopathological features (China).…….……...237

Table S2: Collated expression analysis of DP103 epitope in basal, non-basal subtypes and in the various grades of breast caancer in specimens from Singapore and Canada cohorts of patients.…..……..…….……..…….239

Table S3: Spontaneous metastases from orthotopic mammary fat pad implant with MDA-MB-231-EV and MDA- MB-231-DP103 cells……..…….....241

Table S4: Collated expression pattern analysis of MMP9 and DP103 epitopes in epithelial compartments of ductal specimens from Singapore and Canada cohorts of patients…….....242

Table S5. A) Collated expression analysis of phospho-p65 S276 epitope in the epithelial compartments of ductal specimens.……..……...244
Table S5 B) Collated expression analysis of phospho-p65 S276 epitope and its association with tumor subtype.……..……..……. 244
Table S5 C) Collated expression analysis of phospho-p65 S276 epitope and its association with histological tumor grades……..……244
Table S5 D) Association analysis between expression of DP103 and phospho-p65 (Ser276)……..……..……245

Appendix IV:

Table 1: List of genes that are significantly highly expressed in luminal B than in luminal A breast cancer……..……..…….249
Table 2: list of 19 genes that are supposed to be excellent targets to distinguish between luminal A and luminal B breast cancer……..……..259

List of abbreviations

Abbreviation	Full term
AF domain	Activating function domian
AFP	Alpha fetoprotein
AKT	Serine-threonine protein kinase (Protein kinase B)
ANXA1	Annexin A1
AP-1	Activator protein 1
APC	Adenomatous Polyposis Coli
ASCO	American Society of Clinical Oncology
ATP	Adenosine triphosphate
ATPase6	Adenosine triphosphate Synthase 6
AUC	Area under the curve
BMRS1	Breast Cancer Metastasis Suppressor 1
BRCA1	Breast Cancer 1
BRCA2	Breast Cancer 2
BUS	B-upstream segment
CA-125	Carcinoma antigen 125
CAP	College of American Pathologists
CCNE1	Cyclin E1
CD95	Cluster of differentiation 95
CDC6	Cell Division Cycle 6
cDNA	Complementary DNA
CDT1	Chromatin Licensing And DNA Replication Factor 1
CE mark	*Communauté Européenne mark*
CEA	Carcino-embryonic antigen
CEP17	Chromosome enumeration probes 17
ChIP	Chromatin immunoprecipitation
CHUM	*Centre Hospitalier de l'Université de Montréal*

CI	Confidence interval
CK 5/6	Cytokeratin 5/6
CK 8	Cytokeratin 8
CK18	Cytokeratin 18
CK19	Cytokeratin 19
CLIA	Clinical Laboratory Improvement Amendments
CMA	Cell Microarray
CNV	Copy number variation
DBD	DNA binding domain
DCIS	Ductal carcinoma *in situ*
DFS	Disease-free survival
DMEM	Dulbecco's modified Eagle's medium Nutrient Mixture
DNA	Deoxyribonucleic acid
DP	Digital pathology
DP103	DEAD-Box Protein, 103kD
DRFS	Distant relapse free survival
DUSP2	Dual Specificity Phosphatase 2
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EMT	Epithelial-Mesenchymal Transition
ER	Estrogen Receptor
ErbB1	V-Erb-B2 Avian Erythroblastic Leukemia Viral Oncogene Homolog 1
ErbB2	V-Erb-B2 Avian Erythroblastic Leukemia Viral Oncogene Homolog 2
ErbB3	V-Erb-B2 Avian Erythroblastic Leukemia Viral Oncogene Homolog 3
ErbB4	V-Erb-B2 Avian Erythroblastic Leukemia Viral Oncogene Homolog 4
ERE	Estrogen response element
ERK	Extracellular signal-regulated kinases

ERα	Estrogen Receptor alpha
ERβ	Estrogen Receptor beta
ESMO	European Society for Medical Oncology
ESR1	Estrogen Receptor 1
ESR2	Estrogen Receptor 1
FAS	Fas Cell Surface Death Receptor
FBS	Fetal Bovine Serum
FDA	Food and Drug Administration
FFPE	Formalin-Fixed, Paraffin-Embedded
FISH	Fluorescent in situ hybridization
FKBP52	FK506-Binding Protein, 52 KDa
FOXA1	Forkhead Box Protein A1
GATA3	GATA Binding Protein 3
GPs	G proteins
GSTP1	Glutathione S-Transferase Pi 1
H&E	Haematoxylin and Eosin
HAT	Histone acetyltransferase
HCG	Human chorionic gonadotropin
HER1	Human Epidermal Growth Factor Receptor 1
HER2	Human Epidermal Growth Factor Receptor 2
HER3	Human Epidermal Growth Factor Receptor 3
HER4	Human Epidermal Growth Factor Receptor 4
HR	Hazard ratio
HSP70	Heat shock protein 70
HSP90	Heat shock protein 90
ICC	Intraclass Correlation Coefficient
IDC-NST	Invasive ductal carcinoma, no special type
IFN	Interferon
IGF-1	Insulin-like growth factor 1
IGF-1R	Insulin-like growth factor 1 receptor
IgG	Immunoglobulin G

IHC	Immunohistochemistry
IL-1	Interleukin 1
ISH	*In situ* hybridization
KDa	Kilodalton
Ki-67	Marker of proliferation Ki-67
KLK4	Kallikrein-Related Peptidase 4
KLK5	Kallikrein-Related Peptidase 5
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
LBD	Ligand-binding domain
LCIS	Lobular carcinoma *in situ*
LOH	Loss of hetrozygosity
LYN	Lck/Yes-Related Novel Protein Tyrosine Kinase
MAPK	Mitogen-activated protein kinase
MCM2	Minichromosome maintenance complex component 2
MCM7	Minichromosome maintenance complex component 7
MINDACT	Microarray In Node-Negative and 1 to 3 positive lymph node Disease may Avoid Chemotherapy trail
miRNAs	MicroRNAs
MMP-1	Matrix metalloproteinase-1
MMP-13	Matrix metalloproteinase-13
MMP-8	Matrix metalloproteinase-8
MMP-9	Matrix metalloproteinase-9
MMPs	Matrix metalloproteinases
mRNA	Messenger Ribonucleic Acid
MSN	Moesin
mTOR	Mechanistic Target Of Rapamycin
MYB	V-Myb Avian Myeloblastosis Viral Oncogene
MYC	V-Myc Avian Myelocytomatosis Viral Oncogene Homolog
NCCN	National Comprehensive Cancer Network
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated

	B cells
NIH	National Institutes of Health
NSE	Neuron-specific enolase
OR	Odds ratio
ORC	Origin recognition complex
OS	Overall survival
PAF	Platelet-Activating Factor
PAM50	Prediction Analysis of Microarray 50
PARP1	Poly (ADP-ribose) polymerase 1
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase chain reaction
PI3K	Phosphatidylinositol-4,5-bisphosphate 3-kinase
PKC	Protein kinase C
PLC	Phospholipase C
PR	Progesterone receptor
PRE	Progesterone receptor element
Pre-RC	Pre-replication complex
PRKCA	Protein kinase C, alpha
PSA	Prostate specific antigen
PTEN	Phosphatase and tensin homolog
QS	Quantity score
RAR-β	Retinoic acid receptor beta
RASSF1A	Ras association domain family 1 isoform A
RB1	Retinoblastoma 1
RE	Response element
RNA	Ribonucleic acid
ROC	The Receiver-Operating Characteristic
ROI	Region of interest
ROR	Risk of recurrence
RR	Relative risk
RS	Recurrence score

RT-PCR	Reverse transcriptase–polymerase chain reaction
RTU	Ready-to-use
SAR	Survival after relapse
SBR-EE	Scarff-Bloom-Richardson-Ellis-Elston
sCC1	Citrate buffer
SERMs	Selective estrogen receptor modulators
SERPINA3	Serpin Peptidase Inhibitor, Clade A, Member 3
SET	index of sensitivity to endocrine therapy
SMA	Smooth muscle actin
SNP	Single nucleotide polymorphisms
Sp-1	Specificity protein 1
Src	V-Src Avian Sarcoma
STAT1	Signal Transducer And Activator Of Transcription 1
TAILORx	Trial Assigning Individualized Options for Treatment (Rx)
TCGA	The Cancer Genome Atlas
TDLUs	Terminal duct lobular units
TFs	Transcription factor
TGF-β	Transforming Growth Factor beta
TIMPs	Tissue inhibitors of metalloproteinases
TMA	Tissue Microarray
TNFα	Tumor necrosis factor alpha
TNM stage	Tumor size, lymph node involvement, metastasis
TP53	Tumor Protein P53
TS	Total score
UCSC	The University of California Santa Cruz
UL	Ubiquitin ligase
VEGF	Vascular endothelial growth factor
Vim	Vimentin
WHO	World Health Organization
XBP1	X-Box Binding Protein 1

Remerciements / Acknowledgment

I would firstly like to thank my supervisor; Dr. Louis Gaboury who has directed and encouraged me throughout my Ph.D. His motivation and the fruitful discussions and advices allowed the completion of this work. I am extremely grateful for your encouragement and trust that you had in me

I would like to thank my colleagues of the Histology and molecular pathology platform for their help and company during the past five years. I want to thank them for all our good time, collaborations and also fun. You have given me a pleasant, stimulating and caring environment to work. In particular, I want to thank Julie Hinsinger for her support at the beginning of my Ph.D, for sharing knowledge and assistance in different aspects during my work and also for her help with Visiomorph©, Tissuemorph© Digital Pathology software. Special thanks goes to Mrs. Micheline Fortin, a histopathology technician for preparing the tissue microarray sections. Also, I would like to thank Romain for his help with the bio banking facility. I am also extremely grateful to all my colleagues; Mirella, Melina and Melanie.

A big thanks for Dr. Sylvie Mader, a principal investigator at IRIC; she provided valuable insights and ideas that enriched and improved the clarity of the presentation of our articles. Also I want to thank David Laperriere, a member of Dr. Mader lab, a very kind and talented individual who was always there to help. I am very appreciative of his assistance with the *in silico* analysis.

Dr. Yves St-Pierre, principal investigator at Institut Armand-Frappier; he always focused my attention on the key points in our work. Thanks for your positive criticism that help a lot in improving the presentation of our articles. I deeply appreciate his willingness to answer all my questions, his enthusiastic support of my work and giving me time to review my first two published articles. I am also grateful to Dr. Muhammad Ramzan-Tahir who helped me a lot with the statistics in all our work.

I would like to acknowledge the contributions of the various agencies that have helped finance my doctoral project, the Institute for Research in Immunology and Cancer

(IRIC), the Faculty of Graduate and Postdoctoral Studies of Université de Montreal and The Egyptian Supreme Council of Universities.

Finally, special thanks go to my parents for all the encouragement over the years. I am profoundly grateful to my husband, Medhat, who always encourage me in difficult times and celebrate moments of joy. I could not have done this without him. Thank you for your encouragement, support and your presence.

CHAPTER I

Introduction

1. Hallmarks of cancer

During the last decade, bio-medical research has focused on better understanding the molecular events that play a role in the initiation and progression of cancer. In 2000, Hanahan and Weinberg proposed six hallmarks of cancer in an effort to didactically pinpoint the most salient features of cancer [5]. The hallmarks of cancer are mechanistic alterations in the cell concurring to bring about transformation of normal cells into malignant cells. These hallmarks include: 1) self-sufficiency in growth signals; 2) resistance to antigrowth signals; 3) escape of apoptosis; 4) limitless replicative potential; 5) sustained angiogenesis; 6) tissue invasion and metastasis [5]. In 2011, following new advances in cancer research, Hanahan and Weinberg have proposed two more additional hallmarks: reprogramming of energy metabolism and evasion of immune destruction. They also added two enabling characteristics: genome instability and tumor-promoting inflammation (Figure 1.1) [6].

1) **Self-sufficiency in growth signals:** Normal cells are critically dependent on growth signals to leave G naught and enter an actively proliferating stage. In contrast, cancer cells seem to acquire sustained proliferative capacity through different mechanisms including: autonomous growth factors production by tumor cells, paracrine interactions with surrounding stromal cells, up-regulation of receptors present at the cell surface and structural alterations of the receptor such as receptor truncation leading to unregulated cell proliferation. In addition, constitutive activation of the signaling molecules downstream to the receptors as well as interference with normal mechanisms that attenuate proliferation may also stimulate cancer cell proliferation.
2) **Insensitivity to antigrowth signals:** To maintain normal cellular quiescence and tissue homeostasis, many anti-proliferative signals are required. Signals associated with the cell cycle clock and soluble growth inhibitors like transforming growth

factor-β (TGF-β) are thought to be important in this regard. Many cancer cells are able to inhibit these anti-proliferative signals to induce uncontrolled proliferation.

3) **Evading apoptosis:** Signals that induce apoptosis in normal cells can be generated from the cell surface receptors that bind survival and death factors (e.g. FAS ligand and its receptor CD95) and monitor the extracellular environment. Alternatively, pro-apoptotic signals may emerge from inside the cell via DNA damage, lack of sufficient survival signals and hypoxia. It is now believed that all cancer cells harbor alterations that enable them to escape apoptosis in response to anti-apoptotic regulators and survival factors or, alternatively, decrease production of pro-apoptotic proteins.
4) **Limitless replicative potential:** Cancer cells have the ability to replicate indefinitely and escape senescence. There are different mechanisms that help cancer cell becoming immortal including genetic and epigenetic changes that disable the cell cycle checkpoint control or alter the maintenance of telomeres at length that prevents senescence either through up-regulation of telomerase activity or by recombination-based lengthening of telomeres.
5) **Sustained angiogenesis:** Cancer cell require abundant supply of oxygen and nutrients to grow and expand their tissue mass. By the same token, they also need to get rid of their waste products. These crucial requirements of cancer cells are met by stimulating sustained formation of new blood vessels that subsequently help maintain tumor growth. Mechanisms that foster increased angiogenic ability of malignant cells include either an increase in pro-angiogenic factors such as vascular endothelial growth factor (VEGF) or a loss of anti-angiogenic factors such as endostatin. Recently, pericytes and bone marrow derived stromal cells have been shown to be important components of the tumor microenvironment that have the ability to stimulate cancer related angiogenesis.
6) **Tissue invasion and metastasis:** During the development of a malignant tumor, some pioneering cancer cells escape their primary location to invade the adjacent tissue and travel to distant sites in order to develop new colonies of cancer cells. Invasion and metastasis are complex multi-step processes that involve many cellular changes such as alteration in cell-cell adhesion molecules (E-cadherin and integrin)

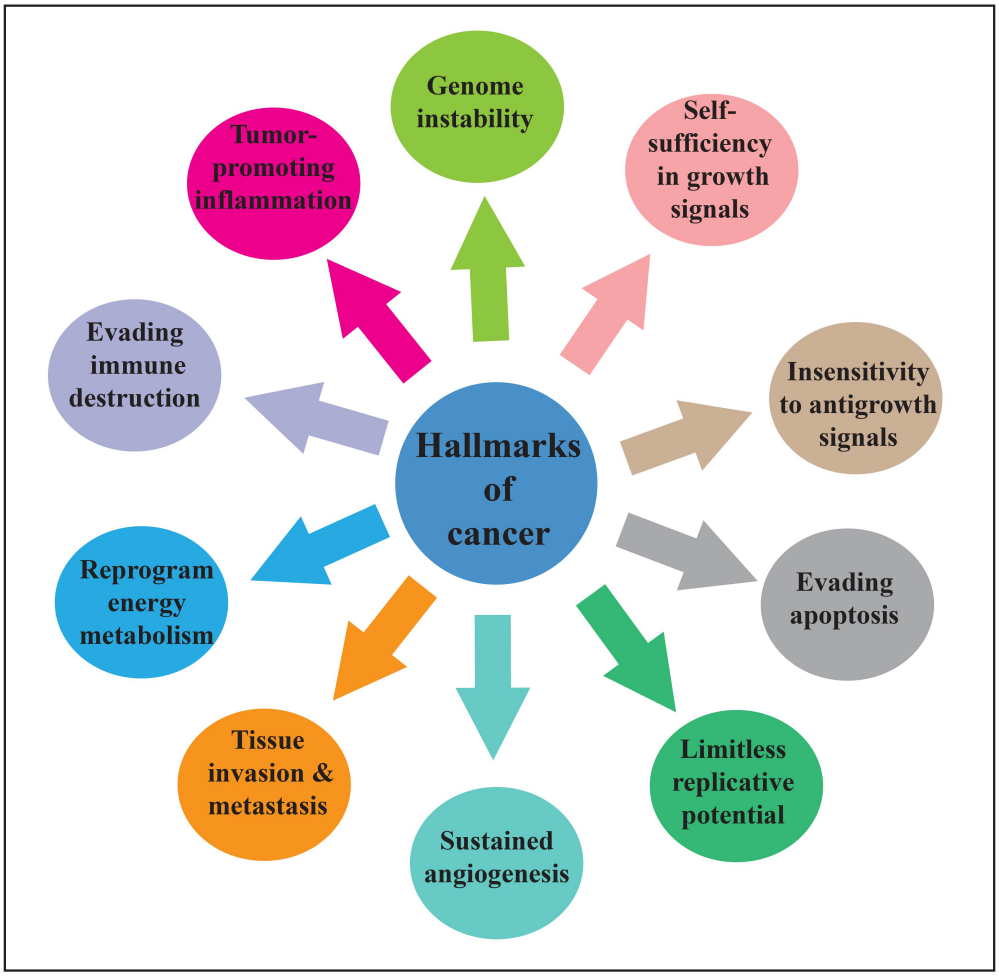


Figure 1.1: Hallmarks and enabling characteristics of cancer.

This illustration includes the eight hallmarks and two enabling characteristics that have been proposed by Hanahan and Weinberg. The hallmarks include: self-sufficiency in growth signals, insensitivity to antigrowth signals, evading apoptosis, limitless replicative potential, sustained angiogenesis, tissue invasion and metastasis, reprogram energy metabolism and evading immune destruction [5, 6].

and increased production of extracellular proteases (matrix metalloproteinases (MMPs)).

7) **Reprogramming of energy metabolism:** Under normal conditions, cells use oxygen for energy production. This process is known as aerobic respiration. However, normal cells can switch to anaerobic respiration when oxygen becomes deficient. In both cases, energy is obtained from breaking down glucose to produce adenosine triphosphate (ATP) that serves as a primary source of energy in the cell [7]. Modifying or reprogramming energy metabolism is an adaptive mechanism of cancer cells to satisfy their energetic needs and fulfill important cell processes such as cell growth and division. Increased glucose transport to cancer cells through up-regulation of glucose transporters is one strategy used by cancer cells to increase their energy production.
8) **Evading immune destruction:** There is an increasing body of evidence to suggest that the immune system acts as a barrier against tumor development and progression. Cancer cells are able to evade immune recognition by the cells of the immune system or by recruiting immune-suppressive inflammatory cells as T-regulatory cells.

Genomic instability and inflammation are considered to be two enabling characteristics that contribute to the occurrence of the above-mentioned hallmarks of cancer [6]. On the one hand, multistep carcinogenesis is depending on the stepwise accumulation of mutations. Indeed, breakdown of one or more components of the genomic maintenance machinery involved in the detection and repair of DNA damages, is one of the principal inducers of genomic instability in cancer cells. On the other hand, inflammatory cells can also contribute to initiation of multiple hallmarks of cancer in several important ways like supplying bioactive molecules to the tumor microenvironment such as growth factors, survival factors and pro-angiogenic factors. Of note, inflammatory cells may deliver mutagenic chemicals such as reactive oxygen species in the vicinity of cancer cells therefore contributing to genetic progression towards a more malignant phenotype.

2. Normal breast tissue

The breasts or mammary glands are paired organs present in both sexes that become fully functional only in nursing females. Mammary glands reach their full development only during the end of pregnancy [8]. Mammary glands are composed of a complex assembly of tissues overlying the chest pectoral muscles [9]. Histologically, one recognizes two main components: glandular tissue and scaffolding stroma. The supporting stromal tissue is made up of mesenchymal-derived cells such as fibroblasts, adipocytes, immune cells, and extracellular matrix [10].

The glandular component of each breast consists of a dichotomously branching ductal–lobular system. Each lobule is made up of acini that form the functional secretory units of the mammary gland, also known as terminal duct lobular units (TDLUs) [11]. In turn, lobules drain into a complex intertwined network of collecting ducts. Normally, mammary ducts and lobules are lined with two cell types: an inner secretory or luminal cells layer and an outer or basal cell layer. The inner luminal cells participate in the production and secretion of milk whereas the outer contractile myoepithelial cell layer facilitates excretion of colostrum and milk into the ducts during lactation [12]. The basal layer includes terminally differentiated myoepithelial cells, mammary stem cells and their progenitors (Figure 1.2) [13, 14]. Moreover, myoepithelial cells are responsible for creating and maintaining the surrounding basement membrane, which is formed of extremely dense layer of collagen IV, laminin and proteoglycans. It is responsible of separating the basal cell layer from the surrounding stroma [15].

Paracrine interactions that occur between myoepithelial and luminal epithelial cells are thought to be important in maintaining several important functions including regulation of cell cycle progression, establishment of epithelial cell polarity, and inhibition of cell migration and invasion [16]. Until recently, it was believed that the great majority of human breast cancers arise from luminal epithelial cells [17]. However, recent molecular studies have raised the possibility that different subtypes of breast cancers may originate from different cellular types such as luminal, myoepithelial and stem cells [18].

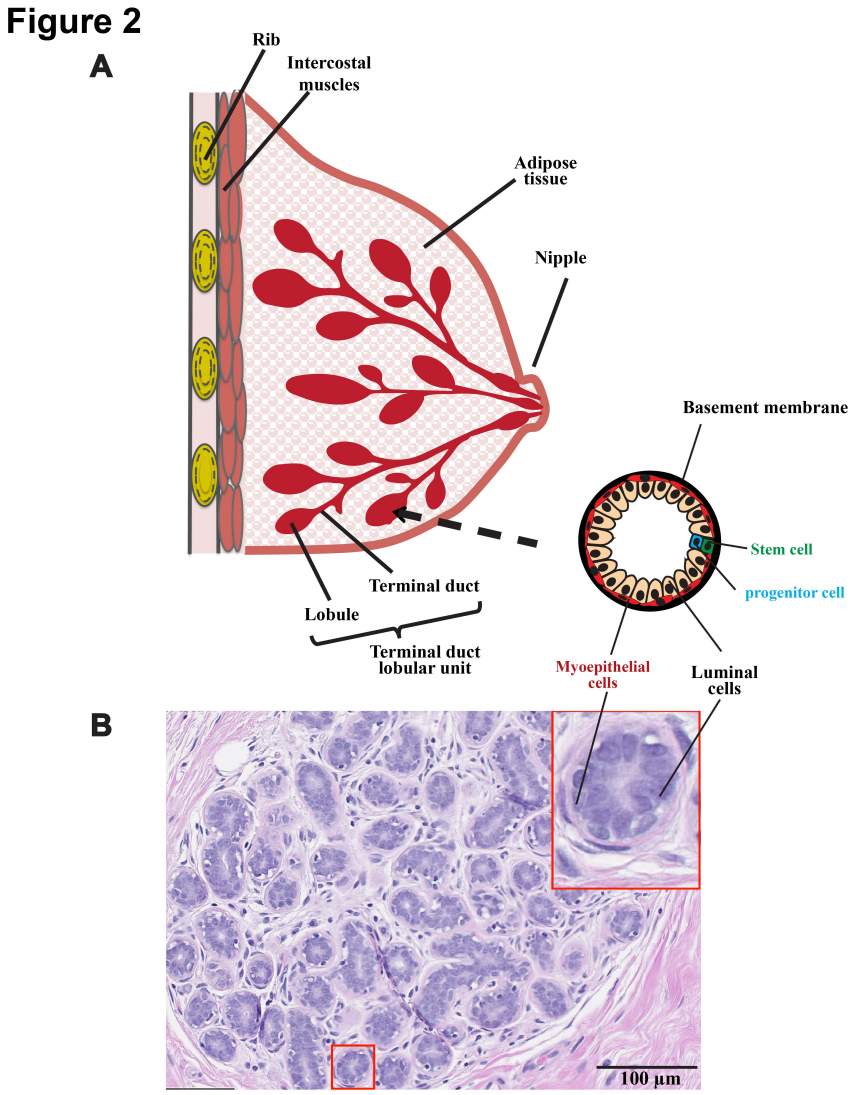


Figure 1.2: Structure of normal breast

A) Diagram depicting the structure of the normal breast. The glandular portion of the breast is composed of a complex branching ductal–lobular system. Each lobule is made up of acini that form the functional terminal duct lobular units (TDLUs). Normal ducts and lobules are lined by an inner secretory luminal cell layer and by an outer layer of basal cells comprising differentiated myoepithelial cells, mammary stem cells and allied progenitors. B) Normal breast lobule stained with H&E showing both glandular tissue and supporting stroma. The inset depicts a normal breast duct lined with an inner luminal cell layer and outer myoepithelial cell layer (Magnification: 20X, inset: 63X).

3. Breast cancer

3.1 Incidence and heterogeneity of breast cancer

Breast cancer is the second most common cancer in women after lung cancer [19]. It is estimated that about 1.7 million new female breast cancer cases are diagnosed every year [20]. According to the 2015 report of the American Cancer Society, 231,840 new cases of invasive breast cancer are diagnosed among women in the United States. An estimated 40,290 breast cancer deaths in women are expected during the same year [21]. This high mortality rate stems mainly from metastasis due to spread of cancer cells to distant sites such as liver, lungs, bones and brain [22].

Breast cancer is a complex and heterogeneous disease. Not surprisingly, risk factors, clinical presentation, pathological features, molecular characteristics and response to treatment vary a great deal from one tumor to the other [23]. The classification of breast cancer which is currently based on clinical and pathological features attempts to classify breast cancer into distinct categories with different prognosis and clinical outcome [24]. Unfortunately, tumors with similar histological features may display divergent clinical behaviors [25]. Recent advances in cancer research based on DNA microarray, functional genomics and gene expression profiling have made important headways in understanding the heterogeneity and complexity of breast cancer [26]. The current molecular classification suggest that breast cancers can be divided into five distinct subtypes; luminal A, luminal B, HER2-positive, basal-like and normal breast-like breast cancer. This approach has allowed the development of prognostic and predictive gene expression signatures and identified new therapeutic targets [27].

3.2 Histological classification of breast cancer

3.2.1 Histological subtypes

The traditional histological classification of breast cancer takes advantage of morphological changes to divide tumors into categories with different prognosis and clinical behavior [24]. Breast cancers are roughly divided into *in situ* carcinoma or invasive breast cancer. Ductal carcinoma *in situ* (DCIS) and lobular carcinoma *in situ*

(LCIS) are considered to be non-invasive or pre-invasive breast cancers involving abnormally proliferating epithelial cells whose expansion in the TDLUs is limited by the basement membrane of the breast ductal or lobular system [28, 29] (Figure 1.3A).

Invasive breast carcinoma results from the spread of cancer cells beyond the basement membrane. In turn, invasive breast cancer can be further classified into ductal carcinoma, no special type (IDC-NST) when it fails to exhibit specific tumor characteristics that allow inclusion into specific histological types such as lobular, tubular, papillary, medullary and mucinous breast cancers (Figure 1.3B-F) [30]. According to the 2012 edition of the World Health Organization (WHO), IDC-NST is considered to be the most common type of breast cancer comprising 40%-75% of all breast cancers leaving only 20%-25% of breast cancers to other special types [31]. Evidently, this histological classification has important limitations since there is an important variability in the outcome of patients belonging to the same histological subtype of breast cancer. Furthermore, left alone (i.e. without additional markers) this classification is not able to guide treatment of breast cancer patients [30].

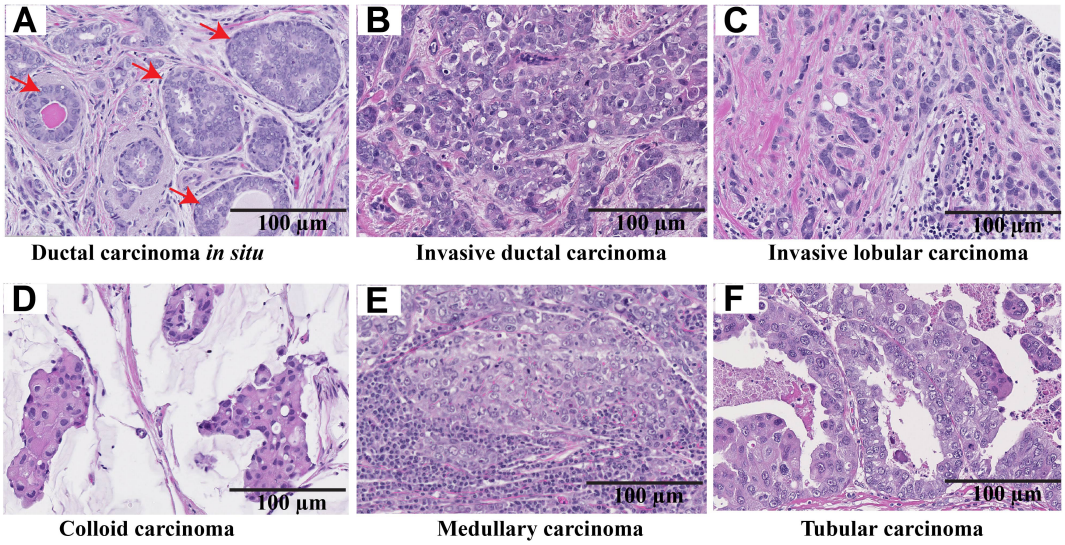


Figure 1.3: Different histological subtypes of breast cancer

A) Ductal carcinoma *in situ*, B) invasive ductal carcinoma, C) Invasive lobular carcinoma, D) Colloid carcinoma, E) Medullary carcinoma, F) Tubular carcinoma

3.2.2 Histological grades

Histological grade of tumors is based on three parameters: degree of differentiation (i.e. extent of tubule formation), proliferation rate (i.e. mitotic index) and nuclear polymorphism (Table 1.1) [24]. Breast cancer patients are ascribed to one of the three different histological grades (Grade I, II & III). Grade I tumors (well differentiated) demonstrate conspicuous tubule formation (> 70%), low mitotic count and only a mild degree of nuclear polymorphism. Grade II tumors (moderately differentiated) have a lower degree of tubule formation (10-70%), intermediate mitotic count and higher degree of nuclear polymorphism. Grade III tumor (poorly differentiated) shows no or very little tubule formation (< 10%), frequent mitoses and a marked degree of cellular polymorphism (Figure 1.4A-C).

Histological grading of breast cancer is a relatively simple and cost-effective method. Genome-wide microarray-based expression profiling studies have provided further evidence that the biological features captured by the histological grade are important to determine tumor behavior [32]. Interestingly, the breast cancer grading system has been incorporated into the criteria used by the adjuvant! Online predictive tool and the 2013 St-Gallen's expert consensus guidelines for adjuvant chemotherapy of breast cancer [33, 34].

3.3 Molecular classification of breast cancer

The rapidly expanding amount of molecular data has led to a better understanding of the heterogeneity of breast cancers and allowed to propose new etiologic pathways leading to breast cancer development [35]. In 2000, gene expression profiling was first developed by the Stanford group in an attempt to classify breast cancer into molecularly distinct groups. Using complementary DNA (cDNA) from 65 samples of breast cancer they sought to determine the levels of expression of 8,102 human genes. Unsupervised hierarchical clustering on these genes revealed "molecular portraits" that give indications about similarities and differences among tumors, and, in many cases, pointed to biological hypotheses [36]. This seminal study was followed by many different attempts

Table 1.1: Parameters of histological grading of invasive breast carcinoma [24].

Score	**1**	**2**	**3**
Tubular score	> 70%	10-70%	< 10%
Mitotic count/10HPF	0-9	10-19	> 20
Nuclear score	Small, regular, uniform	Moderate size & Variation	Large & marked nuclear variation
Total score: range 3-9 Grade I: total score is 3-5 (Well differentiated) Grade II: total score is 6-7 (Moderately differentiated) Grade III: total score is 8,9 (Poorly differentiated)			

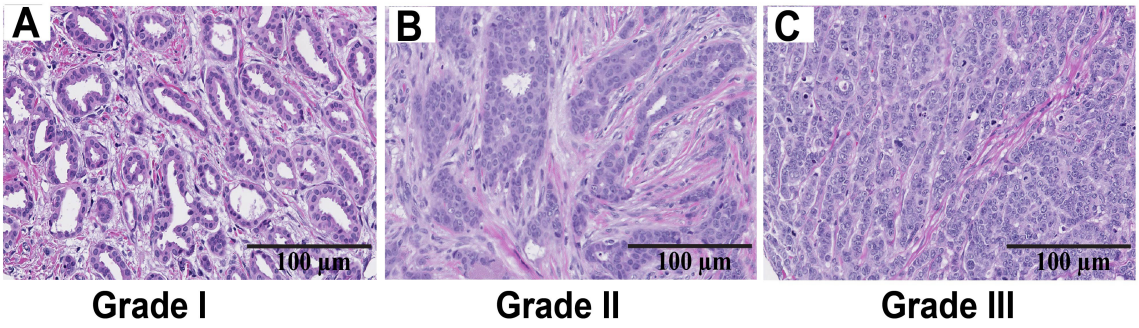


Figure 1.4: Different histological grades of breast cancer

A) Grade I tumors (well differentiated) demonstrate conspicuous tubule formation (> 70%), low mitotic count and only a mild degree of nuclear polymorphism. B) Grade II tumors (moderately differentiated) have a lower degree of tubule formation (10-70%), intermediate mitotic count and higher degree of nuclear polymorphism. C) Grade III tumor (poorly differentiated) shows no or very little tubule formation (< 10%), frequent mitoses and a marked degree of cellular polymorphism

by the same group or by other groups to analyze larger number of patients and validate the reproducibility of these molecular subtypes. Additionally, they also thought fit to further correlate subtypes with clinical outcome in these patients [3, 37-40]. This classification of breast cancer provided, to some extent, explanations on how two tumors with identical clinical and pathological features could behave differently.

The main molecular, intrinsic, subtypes comprised: estrogen receptor positive (ER-positive) breast cancer (luminal A and luminal B) and estrogen receptor negative (ER-negative) breast cancer (HER2-enriched, basal-like and normal breast-like breast cancer). Interferon-rich, apocrine and claudin-low molecular subtypes were added at a later date into the ER-negative molecular breast cancer subtypes (Table 1.2) [1-4]. In addition, Sorlie et al. proposed an additional subset of luminal breast cancer called luminal C which could be distinguished from luminal A and B subtypes by over expression of a novel set of genes shared with basal-like and HER2-enriched subtypes but whose function is currently unknown [37]. However, further studies failed to reproduce the luminal C subtype and, therefore the luminal classification now comprises only luminal A and B subtypes [41].

3.3.1 ER-positive breast cancer: Luminal breast cancers

Luminal tumors are the most common molecular subtype of breast cancer and account for approximately 75% of breast cancers [42]. Luminal breast cancer are known to express gene cluster characteristic of luminal epithelial cells such as Estrogen receptor (ER), Progesterone receptor (PR), cytokeratin 8 and 18 (CK8 and 18), GATA3, FOXA1, XBP1 and MYB [36, 43]. In addition, depending on the expression levels of proliferation-related genes and, to a lesser extent, the expression levels of ER and ER-related genes, hormone responsive breast cancer (luminal breast cancer) can be further classified into either luminal A or luminal B breast cancers [23].

Luminal A breast cancer, which is the less aggressive subtype, express the highest level of ER-related genes and the lowest level of proliferation-related genes. In addition, this type of tumor is usually of low histological grade and is associated with a good prognosis [41, 44, 45]. Endocrine therapy alone is the mainstay of systemic therapy recommended for luminal A breast cancer. Generally, agents such as selective estrogen

Table 1.2: Characteristics of different molecular subtypes of breast cancer [1-4].

Molecular subtypes	ER	PR	HER2	Proliferative marker	Characteristic genes	Prognosis
Luminal A	+	+	-	Low	High level of ER-related genes and low proliferation-related genes	Good
Luminal B	+ +	+ +	- +	High Low/High	Low level of ER-related genes and high proliferation-related genes	Intermediate/ Bad
HER2-enriched	-	-	+	High	High expression of HER2 gene and HER2-related genes (pathway, near HER2 locus). Overexpression of proliferation related genes and lower expression of luminal related genes	Bad
Basal-like	-	-	-	High	Keratin 5, keratin 17, integrinb4, laminin, vimentin, epidermal growth factor receptor (EGFR) and markers of high proliferative state	Bad
Normal breast-like	-	-	-	Low	High expression of genes characteristic of basal epithelial cells and adipose cells, and low expression of genes characteristic of luminal epithelial cells	Good
Interferon-rich	-	-	-	High	High expression of IFN- regulated genes such as STAT1. High expression of gene categories relating to “immune response” and “defense response”.	Intermediate/ Bad
Molecular apocrine	-	-	-	High	Increase androgen signaling	Bad
Claudin-low	-	-	-	High	Low expression of genes involved in tight junctions and cell-cell adhesion, including Claudins 3, 4, 7, Occludin, and E-cadherin. Low expression of luminal genes, high expression of lymphocyte and endothelial cell markers.	Bad

receptor modulators (SERMs) like tamoxifen, pure selective regulators of ER like fulvestrant and third-generation aromatase inhibitors (anastrozole, letrozole, or exemestane) are recommended [46].

In contrast, luminal B is a more aggressive breast cancer subtype. As expected, it is associated with a lower level of expression of ER and ER-related genes, weak or low levels of progesterone receptor expression and increased expression of proliferation-related genes [47]. In addition, activation of the growth factor receptor signaling pathways such as IGF-1R and PI3K/AKT/mTOR is commonplace in this subtype of breast cancer [48]. In contrast to luminal A subtype, luminal B breast cancer is associated with a higher histological grade, sustained proliferation rates, and an overall worse prognosis [49]. Interestingly, the overall survival of untreated luminal B breast cancer is similar to that of HER2-enriched and basal-like subtypes [41]. As a result, treatment of luminal B breast cancer generally combines both endocrine therapy and chemotherapy. According to the 2013 St. Gallen's recommendations, addition of chemotherapy is considered only in patients with intermediate or high risk for recurrence [33]. Although as many as 30% of luminal B breast cancers overexpress HER2 (luminal B/HER2-positive), these tumors maintain the same profile of gene expression that is observed in luminal B/HER2-negative subtype. However, luminal B/HER2-positive tumors require a distinct therapeutic approach, with addition of HER2-targeted therapy such as trastuzumab to chemotherapy [50].

3.3.2 ER-negative breast cancer

HER2-enriched breast cancer

HER2 belongs to the receptor tyrosine kinase family that comprises EGFR (ErbB1, HER1), ErbB2 (HER2), ErbB3 (HER3) and ErbB4 (HER4). The HER2 gene is located on chromosome 17 and found to be amplified in several types of cancers including ovary, endometrium, bladder, lung, colon, and head and neck [51]. Activation of HER2 brings about different cellular responses such as cytoskeletal rearrangement, abrogation of apoptosis and increased cell proliferation [52]. Both HER2 gene amplification and activating somatic mutations are associated with HER2-enriched breast cancer [53, 54].

The HER2-enriched subtype of breast cancer comprises approximately of 15-20% of all breast cancers. It is characterized by elevated expression of HER2 gene and other genes associated with the HER2 pathway or genes located in the vicinity of the HER2 locus on chromosome 17. It entails overexpression of proliferation related genes and a lower expression of luminal associated genes [55]. In addition, 75% of HER2-positive tumors display a high histological grade and more than 40% harbor TP53 mutations [56].

Generally, HER2-positive breast cancers have a poor prognosis. However, in the last decade, the use of anti-HER2 therapy such as trastuzumab has improved the survival of HER2-positive patients both in metastatic disease and in early breast cancer [57, 58]. Interestingly, some tumors lose HER2 expression upon treatment with trastuzumab resulting in a significantly worse relapse-free survival than those with tumors that retained HER2 amplification [59]. Clinical trials combining drugs that inhibit HER2 signaling pathway at different levels showed promising results in terms of response rate and tolerability [60, 61].

Basal-like breast cancer

The basal-like breast cancer represents 10-20% of all breast carcinomas [55]. The incidence of this subtype of breast cancer is highest among premenopausal African-American women. It is characterized by larger tumor size, higher histological grade and regional lymph node involvement [62, 63]. Overall, women with basal-like breast cancers have a worse prognosis and shorter relapse-free life expectancy than women with other types of breast cancer [64]. In addition, basal-like breast cancer is associated with higher rates of metastasis to the brain and lungs [65, 66].

Based on gene expression profiling, the term basal-like breast cancer was coined due to the expression of genes that are normally expressed in normal basal and/or myoepithelial cells. The gene expression cluster characteristic of basal cells includes keratin 5, keratin 17, integrin b4, laminin, vimentin, epidermal growth factor receptor (EGFR) and markers of proliferation [36, 67]. In the clinics, immunohistochemistry (IHC) has been used as a surrogate of gene expression profiling to identify this subtype of breast cancer since it is not only easier but also readily applicable to formalin-fixed, paraffin-embedded (FFPE) samples [68].

There are many ways to define basal-like breast cancer using IHC staining; most commonly one looks at the 'triple-negative phenotype', which is characterized by a lack of ER, PR and HER2 expression [22]. Other authors maintain that the basal cytokeratin profile should also be present including expression of basal cytokeratin 5, 14 and/or 17 [69]. Lastly, it has been proposed that the combined expression of five markers best define this subtype of breast cancer; this panel includes negative expression of ER, PR and HER2 and positive expression of CK5/6 and/or EGFR [70]. Of note, one third of triple-negative tumors have a non-basal genomic profile; these tumors generally have an improved prognosis [22].

When basal-like breast cancers are considered, the overall mutation rate is much higher than for luminal A tumors [43]. Accordingly, basal-like breast cancers exhibit numerous genetic alterations including looses of TP53, RB1 and BRCA1 or amplification of CCNE1 and MYC [71]. It has been reported that close to 20% of basal-like breast cancer harbor either germ line or somatic mutations in BRCA1 or BRCA2 genes [43, 72]. In addition, basal-like breast cancers are generally highly proliferative tumors due to defective TP53 and/or retinoblastoma (RB1) proteins [73, 74].

To date, there is no targeted therapy available for treatment of basal breast cancers. Because basal-like breast cancer fail to express either hormonal receptors or HER2, one does not expect clinical response to any known endocrine or targeted therapy such as tamoxifen, aromatase inhibitors or trastuzumab [75]. According to several clinical trials, basal-like breast cancers are more likely to respond to anthracycline or platinium-based neoadjuvant chemotherapy protocols than any other subtypes of breast cancer. However, despite initial good clinical response to chemotherapy, basal-like breast cancer have a poor prognosis due to higher rate of relapse [76]. Recently, new regimens of therapies have been developed to specifically target Poly (ADP-ribose) polymerase 1 (PARP1), EGFR, mitogen-activated protein kinase (MAPK) and the AKT pathways [22, 77]. So far, none of these therapies have yet reached approval by the U.S. Food and Drug Administration (FDA) [77].

Normal-like breast cancer

Normal-like breast cancer subtype is thought to represent about 5-10% of all breast carcinomas [55]. Perou et al. have clustered the normal-like breast cancers with both normal breast tissues and fibroadenomas. The gene expression pattern in these tumors comprises elevated levels of genes expressed in basal epithelial cells and adipose cells, and low levels of expression in genes characteristic of luminal epithelial cells [36]. Most tumors lack expression of ER, PR and HER2, and, accordingly, these tumors should be classified as triple negative. However, since they lack the expression of CK5 and EGFR they are not considered to belong to the basal-like breast cancers subtype [55]. Of note, normal-like tumors showed the lowest degree of expression of proliferation related genes [78].

Previous studies have raised important doubts about the existence of this molecular subtype [3, 25, 79]. Indeed, these tumors may simply correspond to technical artifacts resulting from a high degree of contamination of frozen samples with normal breast tissue and stromal cells [25, 79]. Given that no case of normal-breast like tumors could be detected when neoplastic cells were isolated by micro-dissection, there are reasons to question the validity of this molecular subtype [80].

Interferon-rich subtype

The interferon-rich subtype of breast cancer is an ER-negative subgroup that has been described for the first time by Hu and colleagues [3]. This subtype is characterized by enhanced expression of either interferon-regulated genes such as STAT1 or some other immune response related genes [3]. The IFN gene cluster has been linked to lymph node metastasis and poor prognosis [3, 81]. Interferon-rich breast cancers are highly proliferative when compared to luminal A tumors [82]. In addition, survival rates are comparable to that of luminal B tumors. The prognosis in these tumors is somewhat better than for all other ER-negative breast cancers [82, 83].

Molecular apocrine subtype

The molecular apocrine group of breast cancer combines the concomitant loss of ER expression with up-regulation of androgen signaling and expression of apocrine

histological characteristics. These tumors share some common features with HER2-positive subtype of breast cancer [2, 84]. Banneau et al. suggested that the development of apocrine breast cancers is due to loss of PTEN at the early stages of tumor development [1]. Cells with apocrine features in the mammary gland originate from metaplastic changes of normal ductal cells leading to morphological changes akin to those seen in apocrine glands of the axilla and perineum [85]. The molecular apocrine tumors are considered to be aggressive and have a poor clinical outcome [86]. Indeed, poor prognostic factors such as high histological grades, increased proliferation rate, lymphovascular invasion and lymph node involvement all accompany apocrine cancers [87].

Claudin-low molecular subtype

In 2007, claudin-low molecular subtypes was first identified and characterized by Herschkowitz and his colleagues using 13 different murine models and comparing the results to that of human breast cancer [88]. They demonstrated that this molecular subtypes is characterized by a lower expression of genes involved in tight-junctions and cell-cell adhesion molecules, including claudins 3, 4, 7, occludin, and E-cadherin. These tumors display low levels of luminal related genes, inconsistent basal gene expression, and a high expression of lymphocyte and endothelial cell markers [88]. Furthermore, there is now compelling evidence to suggest that claudin-low tumors are enriched in cells undergoing epithelial-to-mesenchymal transition (EMT) and showing immune system responses and stem cell-associated features [4, 89].

It has been estimated that approximately 75-80% of these tumors are triple-negative (ER-/PR-/HER2-), while the remaining 15-25% express hormone receptors [90]. Clinically, this molecular subtype of breast cancer appears to be a strong predictor of relapse and poor prognosis [91, 92]. The response rate of claudin-low tumors to standard preoperative neoadjuvant chemotherapy is intermediate between that of basal-like and luminal tumors [4].

3.4 Immunohistochemical markers as surrogate for molecular subtyping and its pitfalls

There is no doubt that gene expression profiling resulted in a substantial advance in our understanding of breast cancer heterogeneity. It also paved the way to molecular classifications of breast cancer with delineation of clinically relevant subtypes. However, there are important limitations that need to be resolved before considering gene expression profiling ready for prime time clinical usage. The reliability and the reproducibility of this technique is questionable as different studies revealed large variations in the expression profiles between cDNA microarrays, difficulties in distinguishing between two similar genes and their spliced variants and inconsistent sequence fidelity of spotted microarrays with only 62%-80% of genes immobilized on the microarray matching the correct sequence of the clones [93-97]. Notably, using different commercially available microarray platforms with identical RNA preparations one observes considerable divergence in the gene expression measurements across different platforms [98]. There are also complexity issues surrounding microarray analysis such as the unyielding nature of data generated by bioinformatics analysis and the cryptic methods used for interpretation [97]. Finally, the high cost of this analysis in comparison to already available pathological approaches is another important issue [99].

Due to the above shortcomings, efforts have been made to compare and contrast the results collected from gene expression profiling of breast tissue with currently existing IHC-based assays using different prognostic and predictive biomarkers [100-102]. IHC is a widely available and relatively inexpensive technique to measure biomarker expression on fresh and FFPE tissues [103-105]. New tools such as automated IHC, whole slide cell scanning and advances in digital pathology softwares to assist IHC scoring are now becoming widely available [106]. Taken together, these factors along with the availability of tissue microarray (TMA) allow using high throughput approach to be used simultaneously on a large number of tissue samples [107]. Importantly, because IHC maintains tissue architecture, one can evaluate tumor and stromal cells separately which is not yet possible with gene expression profiling [99]. To be fair, many difficulties and limitations are also associated with IHC. These include pre-analytical and analytical

factors including specimen fixation, tissue processing, antigen retrieval and the use of revealing reagents. Other analytical and post-analytical issues should also be considered in the selection of antibodies, sensitivity of reagents, choice of antibody type and clone and interpretation of the results [108, 109]. Standardization of various aspects of IHC techniques needs to be resolved before producing sound clinical results.

It has been suggested that IHC can be used clinically to define biologically relevant breast cancer subgroups comparable to those identified by gene expression profiling. Nielsen et al. have demonstrated that a panel of four antibodies (ER, HER1, HER2, and cytokeratin 5/6) can identify basal-like tumors [110]. Others have defined basal-like breast cancer using a panel of five biomarkers (ER, PR, HER2, EGFR, cytokeratin 5/6) that presumably allow predicting more accurately breast cancer survival [70]. In 2011, the 12th St. Gallen's international breast cancer expert panel approved a new classification of patients for therapeutic purposes. It has been suggested that using ER, PR, HER could distinguish molecular subtypes and adding Ki-67 with 14% threshold could help separate luminal A from luminal B subtypes [111]. However, both Guiu et al. and Lips et al. claimed that using Ki-67 with a 14% index has no predictive or prognostic value in luminal breast cancer and rather recommended the use of IHC for ER and HER2 alone for the detection of clinically relevant subtypes of breast cancer [112, 113]. Newer recommendations presented at the 13th St Gallen's international breast cancer expert panel in 2013 proposed to change the Ki-67 index from 14% to 20% to better distinguish luminal A from luminal B breast cancer [45].

4. Cancer biomarkers

According to the American Cancer Society, a tumor marker is defined as a molecule or process that refers to the presence of cancer [114]. Tumor markers are often produced by cancer cells themselves or by other tissues in response to the presence of cancer cells or other associated conditions such as inflammation. Tumor biomarker detection can rely on non-invasive approaches when it is used on body fluids such as blood and urine. Alternatively, tumor markers can be detected on a biopsy of a solid tumor.

An ideal tumor marker should clearly distinguish normal and cancer cells, be easy to use, reproducible, reliable, sensitive (low false negative rates), specific (low false positive rate) and cost effective. Unfortunately, biomarkers with ideal sensitivity and specificity profiles are difficult to find [115]. Davis et al. rightly emphasized that lowering medical cost together with better patient care should be the main objectives of a new biomarker [116].

4.1 Biomarkers: from discovery to clinics

The essential steps leading to the development of a new biomarker include the following; biomarker discovery, validation and qualification. Two general approaches have been proposed for biomarker discovery. The first one is a stepwise, hypothesis-driven approach that assesses molecules believed to be potential biomarkers [117]. The other is a discovery-based approach, looking at gene-expression patterns or mass-spectroscopic peaks that could identify new candidate genes, proteins or expression signatures [118].

A biomarker should pass through two distinct exhaustive processes from discovery to the point of use that include the analytical validation method and the clinical qualification processes [119].

A) Biomarker validation: this step consists of evaluating the assay, its performance characteristics, and the optimal condition to be used in order to generate reproducible and accurate results [120]. The validation process should be tailored to meet the intended use of the biomarker, also known as a "fit-for-purpose" approach [121]. Pepe et al. has enumerated the following steps in the process of biomarker validation [122]:

1. Preclinical exploratory studies: to recognize and prioritize potentially useful biomarkers.
2. Clinical assay development and validation: to assess the ability of a biomarker to distinguish subjects with cancer from normal individuals.

3. Retrospective longitudinal studies: to study if a biomarker can detect cancer at an early stage before it becomes clinically detectable.
4. Prospective screening studies: to assess biomarker accuracy by determining its detection rate and false referral rate.
5. Cancer control studies: to assess if a biomarker reduces the burden of cancer on the population.

B) Biomarker qualification:

Validated biomarkers could be qualified for clinical use by different approaches [123, 124]:

1. Passive approach: This approach is based on what is already published in the literature to accept a biomarker for clinical use. However, this process is time consuming, inefficient and unreliable for biomarker qualification.
2. Active approach: In this approach, qualification depends on a comprehensive process that provides clearly defined and precise roles to help the development of new biomarkers. A FDA biomarker qualification pilot study is an example [125].

4.2 Classifications of tumor biomarkers:

Tumor biomarkers can be classified in many ways such as the nature of the biomolecules under scrutiny (e.g. DNA, RNA and Proteins biomarkers), tissue distribution (e.g. blood, urine or tissue biomarkers) and potential clinical application (e.g. risk detector, diagnostic, prognostic and predictive biomarkers).

4.2.1 Classifications of tumor biomarkers according to their potential use

Risk detection biomarker

These markers are used to determine whose patient is at risk of developing certain type of cancers [126]. BRCA1 and BRCA2 are good examples of risk detectors that can be applied to high-risk group such as those with a strong family history of breast or ovarian cancer [127].

Screening biomarker

These biomarkers should be able to detect early stages of tumor development before the onset of clinical manifestations. The WHO has proposed specific criteria to consider a specific biomarker to be a good screening test. These criteria include the following: the screened tumor should occur with a high incidence or be associated with a significant mortality rate, the progression of the tumor should be fully characterized, facilities for early diagnosis and treatment especially for those at an early stage of the disease should be available, the test should be acceptable to the population, cost-effective and endowed with a high sensitivity and a high specificity [128]. Notably, the potential benefit of a tumor screening strategy should outweigh the potential harms that could result from the intervention to prove the disease.

After approval by the FDA in 1986, Prostate specific antigen (PSA) became the most widely used biomarker for prostate cancer screening [129]. However, patients with high level of PSA should not be considered to suffer from prostatic cancer but instead should be investigated further. CA-125 is another screening biomarker that has been recommended by American Congress of Obstetricians and Gynecologists for women with average risk of developing ovarian cancer [130]. However, one must bear in mind that benign conditions such as leiomyoma, endometriosis and inflammatory conditions in the abdomen such as peritonitis can cause elevated level of CA-125, implying that this test is not specific for ovarian cancer [131].

Diagnostic biomarker

Diagnostic biomarkers are useful to help detecting certain types of cancers. They may also help in establishing a differential diagnosis during the course of the evaluation of a tissue mass (cancer or benign condition), or in determining the likely site of origin of a tumor (primary or metastasis) [126]. These biomarkers should have high sensitivity and specificity in order to decrease the likelihood of falsely positive (positive results in the absence of the disease) and falsely negative (missed diagnosis of real disease) results [132]. Examples of the most common diagnostic markers include serum level of alpha fetoprotein (AFP) for liver cancer [133] carcino-embryonic antigen (CEA), neuron-specific enolase (NSE) or cytokeratin 19 (CK19) in lung cancer [134].

Prognostic biomarker

Prognostic biomarker help providing information about the behavior of a tumor, regardless of the type of treatment envisaged: aggressiveness, probability of relapse, and risk of developing metastases [126, 135]. Such markers could help deciding the need for further treatment. For example, high level of human kallikreins; KLK4 and KLK5 are associated with aggressive forms of ovarian cancers [136]. Another example is the lack of ER, PR and HER2 in breast cancer suggesting the presence of an aggressive form of breast cancer and a higher risk of developing metastases [64]. HCG and AFP are also considered prognostic markers and have been recommended by The American Society of Clinical Oncology (ASCO) to monitor for recurrence in patients treated for advanced seminoma [137].

Predictive biomarker

Predictive biomarkers can predict the response of a tumor to a given therapy before the initiation of treatment or predict which therapy is likely to be effective [126]. For instance, ER-positive breast cancers may benefit from anti-estrogen such as Tamoxifen or aromatase inhibitors [138]. Breast cancer patients with HER2 gene amplification should receive trastuzumab (Herceptin®), a monoclonal antibody that binds to the HER2 receptor [45]. In addition, it has been reported that Ki-67 is a good predictive marker to guide the use of adjuvant therapy in hormone receptors positive breast cancer [139]. Colorectal cancers with activating mutation of KRAS that predicts resistance to EGFR-specific targeted therapy with cetuximab and panitumumab is another example [140].

4.2.2 Classifications of tumor biomarkers according to the biomolecules

1) Genetic biomarker (DNA biomarker)

Genetic biomarkers include any of the following alterations of DNA: single nucleotide polymorphisms (SNP), loss of hetrozygosity (LOH), copy number variation (CNV) of genes and structural variations in chromosomes such as insertion, deletion, translocation, duplication and inversion [141]. For instance, germ line mutation of BRCA1 and BRCA2 genes could be considered as risk detectors as they help identifying

individuals with high risk of developing breast and ovarian cancer [127]. Another example concerns HER2 gene amplification in breast cancer that is considered to be a marker of poor prognosis and an indicator of a higher risk of developing brain metastasis [142]. Moreover, HER2 overexpression can help predict the clinical benefit to anti-HER2 targeting therapy (Trastuzumab), increased sensitivity to chemotherapy regimens containing taxanes and anthracyclines, and resistance to tamoxifen [143, 144]. Furthermore, somatic mutations of mitochondrial DNA such as ATPase6 gene alterations have been recently suggested to be potential biomarkers of breast cancer [145, 146].

2) Epigenetic biomarker

Epigenetics describes heritable changes in cellular phenotype due to changes in the regulation of gene activity and expression that are independent of alteration in the DNA sequence [147]. Some epigenetic changes are used as tumor markers such as aberrant DNA methylation and elevation of global histone deacetylation [148, 149]. Interestingly, epigenetic modifications that occur in the tumor tissues can also be detected in some biological fluids such as urine and serum implying that it could be useful as non-invasive screening tools [150].

Previous studies have reported the usefulness of DNA methylation as epigenetic marker in breast cancer. For example, methylation analysis of a panel of genes consisting of APC, GSTP1, RAR-β and RASSF1A was shown to help in the detection of breast cancer [151]. It has also been shown that epigenetic silencing of the Breast Cancer Metastasis Suppressor 1 (BMRS1) gene due to methylation of its promoter helps predicting an adverse prognosis in breast cancer [152]. Other studies have reported the use of DNA methylation as a marker to help monitor efficacy of breast cancer treatment such as neoadjuvant chemotherapy and tamoxifen [153, 154].

3) Expression biomarkers (RNA biomarkers)

Detection of differentially expressed genes is an interesting approach to new cancer biomarkers discovery [155]. For example, gene expression profiling signatures are able to predict the clinical outcome in breast cancer patients [156]. Recently, microarray gene expression profiling and real time-PCR have been used successfully in detecting

molecular subtypes of breast cancer. We also witnessed the development of multigene diagnostic tests to predict breast cancer recurrence (Oncotype DX or 21 gene assay) or (PAM50 or 50 genes assay) and metastasis (Mammaprint or 70 gene assay) [99, 157].

MicroRNAs (miRNAs) are short, non-coding RNAs that play an essential role in the regulation of gene expression through pairing to specific mRNA targets leading to their degradation and/or translational inhibition [158]. Although the notion of using miRNA as tumor biomarker has not yet been introduced in clinical practice, many reports have indicated specific patterns of miRNA expression in certain types of cancer. Indeed, it has been shown that the expression of MiR-210 is a strong potential prognostic marker in breast cancer [159, 160]. A recent study has reported the use of a four-microRNA signature as a novel diagnostic and prognostic biomarker in triple negative breast cancers [161]. Furthermore, miRNA can also predict the response to different types of cancer treatment such as elevated levels of miR-125b in breast cancer can predict resistance to taxol-based regimens [162].

4) Protein biomarker

The proteome of cancer cells suffers from many aberrations that result from the translation of accumulated gene defects in the DNA. Proteome is also responsible for determining cancer phenotypes and regulating tumor behavior [163]. Since proteins are the main components of signaling pathways in both normal and cancer cells, they are potentially valuable biomarker molecules [164]. Despite the large number of putative biomarkers, there is some merit to recall that the only FDA-approved biomarkers in clinical use are proteins [165].

Although there are thousands of protein tumor biomarkers identified by proteomic studies, only a few of them have been found to be clinically useful mainly because of insufficient validation studies [164]. The latest guidelines of the ASCO in 2007 and the National Academy of Clinical Biochemistry Laboratory Medicine Practice in 2008 recommended the use of ER, PR and HER2 protein expression in invasive breast cancer to predict response to therapies such as hormonal therapy and Trastuzumab [166, 167].

4.3 Problems associated with biomarkers

Despite the presence of thousands of publications about potential biomarkers including their applications and usefulness, only a handful of them have become clinically approved test. This state of affair plainly indicates that the transfer of biomarkers from the bench to clinical application is a lengthy process filled with lots of pitfalls and limitations [115]. The position of the biomarker in the spectrum between research tool and clinical end point depends on the accuracy and strength of experimental demonstration required to achieve method validation [168]. Of note, most of the potential biomarkers to date fail to show adequate sensitivity and specificity to allow them to be used as clinical tests [169]. Some of the potential limitations in the processes of biomarker research are outlined below.

4.3.1 Specimens' problems:

1. Lack of standardization of sample collection: inconsistency in storage conditions, sample handling and who decide to include a given sample in the study can all affect the levels of biomarker detected or lead to false detection due to artifacts present in the samples [169, 170]. Another set of variables such as the type of fixative, time of hot and cold ischemia, time of fixation, inclusion method, time of paraffin-embedding, thickness of the sections, storage of unstained slides and the methods for antigen retrieval are all major sources of variability when comparing the results of different studies [118]. Stringent quality control during specimen collection, storage and handling is crucial for successful biomarker identification. In addition, specialists such pathologists should pay careful attention to select the right samples and appropriate controls. In particular, care must be to obtain the right quantity (enough tissue) and quality (for example; containing cancer cells not only stroma) of tissue for biobanking [171]. Interestingly Moore and his colleagues reported that using high quality controls during tissue collection is a way to improve the overall quality of the specimens, storage and handling [172].
2. Insufficient number of samples to be included in a study are major problems in the process of biomarker identification [173]. Zolg reported that to confirm the results

obtained from a small number of sample, second and third discovery rounds are required to complement the results of the first round [174]. However, to overcome the problem of sample insufficiency Poste G. suggested using hundreds or even thousands of matched diseased and control samples to satisfy regulatory requirements and to demonstrate the clinical usefulness of the tested biomarkers [175].

3. Unbalanced grouping: using samples from patients and controls belonging to different groups such as different race or ages can greatly affect the outcome of a study [169]. To overcome this limitation, people should be matched for as many variables as possible. However, at least their age, sex, race, ethnicity and lifestyle such as alcohol consumption and smoking should be matched [115, 175].
4. Difference on how the patients/samples are assessed, who assessed these patients/samples and how these assessments are reported can also be critical. Detailed history and information about each patient/sample such as the type and grade of cancer are crucial for a good study design. This detailed data will aid in improving samples stratification (e.g. patients with same grade of cancer are in the same group) that could help in detecting biomarkers especially in slowly progressive diseases [115]. In addition, using universal-standardized reporting methods is essential to help comparing different biomarker trails [115].
5. Single analysis of the samples included in biomarker identification studies could affect the final results. Elias and his coworkers study on large scale proteomic analysis demonstrated that reliable results could be obtained only by analyzing samples multiple times [176]. Furthermore, Drucker & Krapfenbauer suggested that samples should be analyzed in triplicates and, one should improve validity of measurements by calculating the mean and standard deviation [115].
6. Another point to consider is the use of tumor cell lines types [177, 178]. Using cell lines may be easier to handle and more accessible as surrogates for patients' tissues. However, there are many disadvantages associated with cell lines including genotypic and phenotypic drift away from the originating tumor during culture work, intra and inter-laboratory cell line heterogeneity and cell line cross-contamination with other cell lines used in the laboratory [179]. Although using replicates (3-10) of each cell line to create a comprehensive biomarker profile is a good approach, it leads to an

increase in the cost of the study [180, 181]. It is highly recommended that scientists ensure that the proprieties of each cell lines they use during the course of their investigations match those of equivalent cell types in vivo [175].

4.3.2 Personnel and instrumental problems

Physical conditions in the laboratory and inadequate instrumentation, failure to calibrate and validate instruments, variations both within and between equipment can all be limiting factors that affect the final outcome of biomarker studies. For example, it has been reported that proteome analysis using different analytical approaches resulted in a discordant number of proteins detected by each method [182, 183]. Moreover, inadequate training of technicians, biochemists and bioinformatics specialists dealing with clinical specimens, faulty pieces of equipment and improper data algorithms can all lead to flawed results. This is especially true of new technological advances being used in biomarker research, where highly trained personnel need to interpret complex set of data [184]. To overcome these limitations, FDA and NIH have set up recommendations to oversee the quality of techniques used in biomarker research [115]. Moreover, FDA has suggested that central reference laboratories be selected to conduct testing of new biomarkers for one or two years prior to its release to the market [185].

4.3.3 Data analysis and interpretation

Lack of standardization in conducting analytical methods and interpreting data are important issues to be taken into account in the field of biomarker research. Most published studies have categorized biomarkers with reference to positive/negative cut-off points. However, this oversimplification probably reduces the value of any given biomarker, as it does not reflect the complexity of many biological phenomenon [118]. For example, although classification of ER-positive breast cancers into two distinct groups; luminal A and luminal B based on proliferation-related genes has been accepted in breast cancer research [23, 80], contrasting views have emerged suggesting that ER-positive breast cancers may form a continuum of tumors with both good and bad prognosis at each end of the spectrum [186-188].

Appropriate statistical methods are required throughout the entire biomarker development pathway [189]. Lee et al. stated that using quantitative analytical methods that are capable of generating reliable data constitute solid ground for statistical assessment of the predictive value of biomarkers [168]. In addition, criteria on how to develop high-quality data project should be established to help comparing outcomes of biomarker research between different research labs or clinical trial studies [115]. Poste G. has suggested a stringent approach to overcome problems related to data analysis including defining intended algorithms for statistical and computational analysis, handling indeterminate results and outliers to resolve clinical false-positive results, estimating clinical sensitivity and specificity with 95% confidence intervals and assessing positive and negative biomarker predictive power [190].

5. Biomarkers in breast cancer

5.1 Established biomarkers in breast cancer

5.1.1 Estrogen receptor (ER)

Estrogen receptor (ER) is a member of the nuclear hormone receptor superfamily located in the cytoplasm and translocate to the nucleus upon activation [191]. ER has two distinct isoforms; ERα and ERβ that are encoded by different genes ESR1 and ESR2 respectively. Each gene is located on different chromosome (locus 6q and locus 14q, respectively) [192]. Both isoforms show considerable sequence homology, each consisting of six domains as depicted in figure 1.5A [193]. Their tissue distribution is quite different as ERα is expressed in uterus, testis, pituitary, ovary, kidney, epididymis, and adrenal gland while ERβ is found in prostate, ovary, lung, bladder, brain, uterus, and testis [194]. Normal breast tissue showed scattered ER-α staining in the ductal epithelium and only occasional nuclear staining for ER-β [195]. ERα and ERβ have different transcriptional activities and opposing actions on cellular processes such as proliferation, apoptosis and migration [193, 196]. However, in the following section we will focus only of ERα (ER).

ER can exert both genomic and non-genomic effects (Figure 1.5B) [197, 198]. Genomic effects happen through interaction between ER and DNA inside the nucleus, either in a ligand-dependent or ligand-independent fashion. In the presence of a ligand such as estrogen, the receptor forms a complex at the ligand-binding site of ER [199]. This complex then migrates to the nucleus where it forms dimer with other ER or, alternatively with direct DNA-binding transcription factors (TFs). Dimers attach themselves to DNA directly at the estrogen responsive element (ERE) or indirectly through tethering interaction with other transcription factor such as activator protein 1 (AP-1) or specificity protein 1 (Sp-1) [198]. At these sites, ER recruits co-activators and co-repressors that lead to modulation of gene transcription through the recruitment of co-regulatory proteins such as histone acetyltransferase (HAT) and ubiquitin ligase (UL). However, ER can also be activated in a ligand-independent manner through activation of ERK and AKT serine/threonine kinases by different growth factors [197, 199]. In that case, the activated kinases phosphorylate ER leading to its activation, dimerization, binding to DNA and subsequent gene regulation [197, 199]. On the other hand, non-genomic effects occur when the ligand-ER complex, present in the cytoplasm or associated with the plasma membrane activates protein kinase cascades (e.g. SRC, PI3K and G proteins (GPs)), phosphatases or increase ion fluxes across membranes. These effects result in rapid physiological responses independent of gene regulation [199-201].

ER is one of the few biomarkers currently in use in the clinical management of breast cancer patients. In normal breast tissue, only 4-15% of epithelial cells express ER [202]. Low expression of ER was reported in normal breast tissue of women at low risk of developing breast cancer [203]. In contrast, overexpression of ER has been observed in normal breast tissue of breast cancer patients [204]. Taken together, the above mentioned data support the notion that ER can be considered a risk factor detector as its overexpression in normal breast tissue augment the risk of breast cancer development [205]. Furthermore, ER has a prognostic value as most of ER-positive breast cancers fare better than ER-negative tumors [206]. However, not all ER-positive tumors have a good prognosis as untreated luminal B breast cancer, which is ER-positive, has an unfavorable prognosis similar to that of HER2-positive and basal-like subtypes of breast cancer [41].

Functional domains of ER

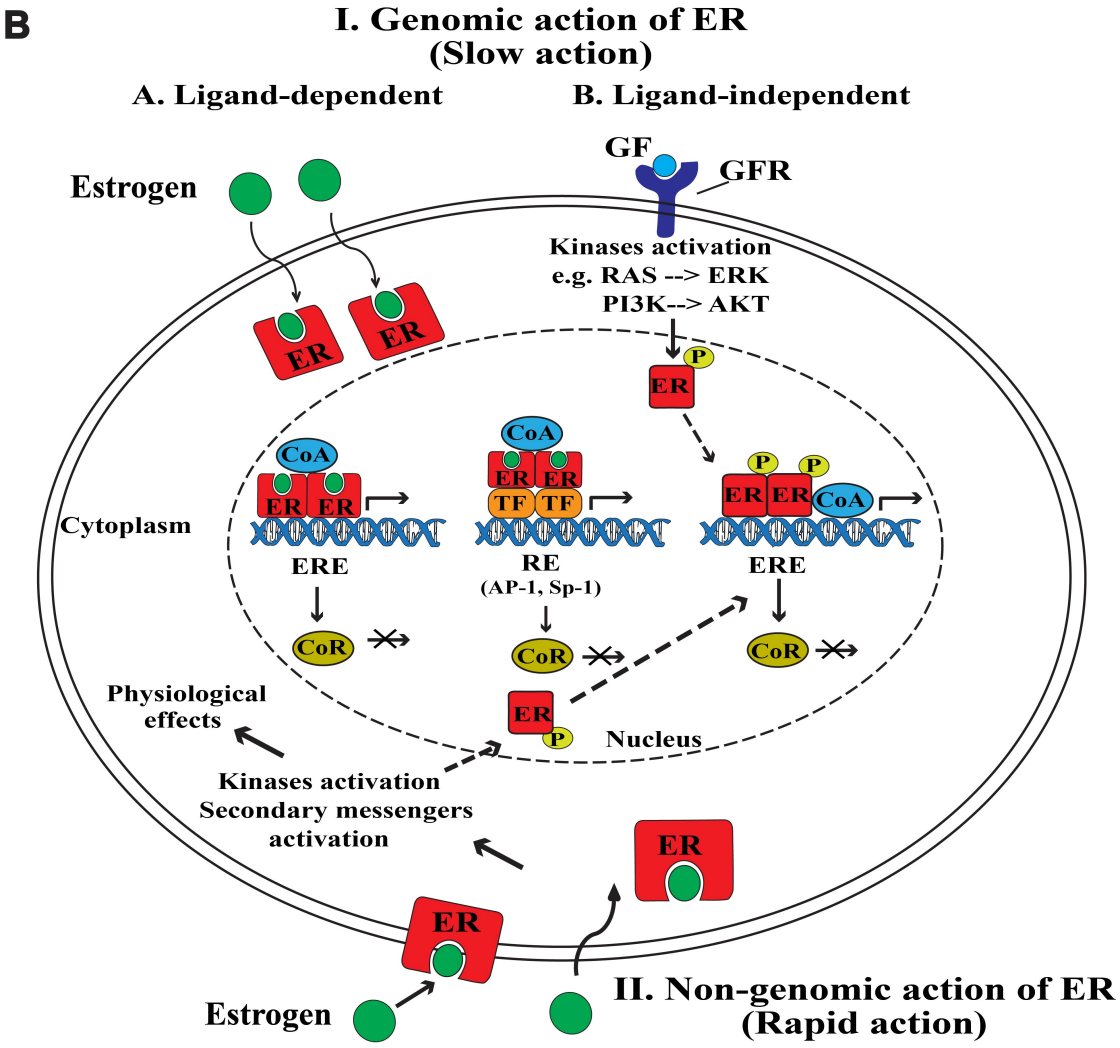


Figure 1.5: Structure and different mechanisms of action of ER

A) ER has the following structural domains: DNA binding domain (DBD), a flexible hinge region that is immediately adjacent to the DNA binding domain, a ligand-binding domain (LBD), two activating function (AF) domains (AF-1 is located on the N-terminal domain and AF-2 which is ligand-dependent). B) Genomic and non-genomic actions of ER: I. Genomic effects happen through interaction between ER and DNA inside the nucleus, either in a ligand-dependent (A) or ligand-independent fashion (B). In the presence of a ligand such as estrogen, the receptor forms a complex at the ligand-binding site of ER. This complex then migrates to the nucleus where it forms dimer with other ER or, alternatively with direct DNA-binding transcription factors (TFs). Dimers attach themselves to DNA directly at the estrogen responsive element (ERE) or indirectly through tethering interaction with other transcription factor such as activator protein 1 (AP-1) or specificity protein 1 (Sp-1). At these sites, ER recruits co-activators and co-repressors that lead to modulation of gene transcription through the recruitment of co-regulatory proteins. However, ER can also be activated in a ligand-independent manner (B) through activation of ERK and AKT serine/threonine kinases by different growth factors. II. Non-genomic effects occur when the ligand-ER complex, present in the cytoplasm or associated with the plasma membrane activates protein kinase cascades (e.g. SRC, PI3K and G proteins (GPs)), phosphatases or increase ion fluxes across membranes. These effects result in rapid physiological responses independent of gene regulation [193, 197, 198]

ER has also a predictive value in breast cancer. Indeed, assessing ER status allows one to predict response to hormonal therapy depending on the level of its expression. Quantitative assessment of ER expression using IHC staining is routinely performed on all breast cancer patients [207]. Allred score stratifies breast cancer patients into two groups according to the probability of responding to hormonal therapy [208]. It is composed of the sum of the proportion (scale of 0-5) and average intensity (scale of 0-3) scores of positive tumor cells resulting in a total score (TS) ranging from 0 to 8 which reflects the staining signal of ER [209]. Tumors with total scores ≥ 3 (corresponding to 1% to 10% of weakly positive cells) are considered to be positive. This cutoff point was detected based on the analysis of disease-free survival (DFS) in a study involving patients receiving any adjuvant endocrine therapy [210]. Tumors with high expression of ER respond better to hormonal therapy while those with low ER status generally benefit from chemotherapy [138, 211].

De novo and acquired resistance to hormonal therapy has been reported in a subgroup of ER-positive breast cancer [212]. Mechanisms involved in resistance to hormone therapy include loss or modification in the ER expression, cross talk between ER and growth factors such as EGFR and HER2, loss of PTEN, constitutive activation of the PI3K/Akt pathway, altered expression of specific microRNAs and genetic polymorphisms involved in tamoxifen metabolic activity [213-216]. Consequently, new treatment strategies are being developed to reverse hormonal resistance and/or amplify the sensitivity of breast cancer cells including the addition of anti-HER2 (Trastuzumab) or chemotherapy to anti-estrogen agents [213, 217].

5.1.2 Progesterone receptor (PR)

Progesterone receptor (PR) is another member of nuclear steroid receptor superfamily that rapidly shuttles between the cytoplasm and the nucleus [218]. It comprises three isoforms, PR-A, PR-B and PR-C with different molecular weight: 94 KDa, 116 KDa and 60 KDa respectively [219, 220]. All isoforms are translated from the PGR gene, which is located on chromosome 11q22 [221]. PRs are composed of the

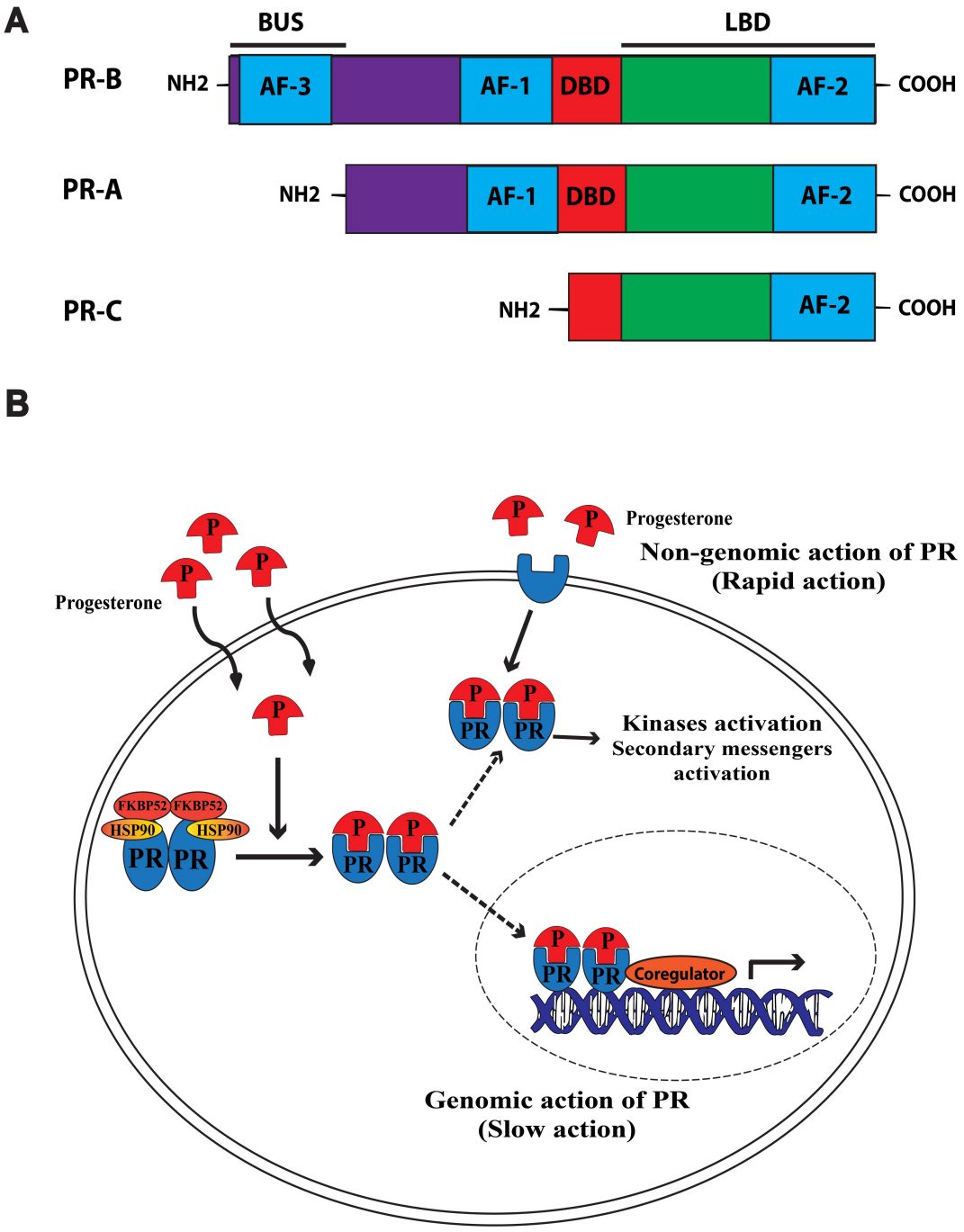

A
BUS
LBD
PR-B
NH2
AF-3
AF-1
DBD
AF-2
COOH
PR-A
NH2
AF-1
DBD
AF-2
COOH
PR-C
NH2
AF-2
COOH
B
Progesterone
Progesterone
Non-genomic action of PR
(Rapid action)
Kinases activation
Secondary messengers
activation
FKBP52
FKBP52
HSP90
HSP90
PR
Coregulator
Genomic action of PR
(Slow action)

Figure 1.6: Structure and mechanism of action of PR

A) This diagram shows the structural domain of the three isoforms of PR, which include PR-A, PR-B and PR-C. PR-A is composed of: DNA binding domain (DBD), a flexible hinge region that assist in DNA binding, a ligand-binding domain (LBD), and multiple activating function (AF) domains required for transcriptional activity. PR-B contains additional unique B-upstream segment (BUS) absent in PR-A. As for PR-C, it has no BUS domain and part of the DBD. B) Genomic and non-genomic actions of PR. In the absence of ligand (progesterone), PR forms a multi-protein complex with chaperon proteins such as heat shock proteins (HSP70 and HSP90) and FKBP52. In the cytoplasm, binding of progesterone to PR stimulates its dissociation from chaperon proteins followed by homo-/hetero-dimerization of PR. These dimers translocate to the nucleus where they bind to the DNA through progesterone receptor element (PRE), AP-1 or SP-1 to induce recruitment of co-regulators (genomic action). However, PR dimers, present in the cytoplasm or associated with the membrane can activate mitogenic protein kinases such as Src and MAPK (non-genomic action) [218, 220, 222, 223]

following structural domains: DNA binding domain (DBD), a flexible hinge region that assist in DNA binding, a ligand-binding domain (LBD), and multiple activating function (AF) domains required for transcriptional activity. PR-B contains additional unique B-upstream segment (BUS) absent in PR-A (Figure 1.6A) [218, 222]. As for PR-C, it is transcriptionally inactive due to the lack of BUS and part of the DBD [220].

PR plays an essential role in normal human female reproductive tissues such as uterus and mammary gland and in non-reproductive tissues such as bone, brain and cardiovascular system [223-226]. It has been reported that PR-A and PR-B are equally expressed in most human target cells suggesting involvement of alternative mechanisms to explain the diversity of progesterone actions [227]. Progesterone receptor shares common features with ER. In that it also functions through genomic and non-genomic signaling pathways (Figure 1.6B) [222, 223]. In the absence of ligand (progesterone), PR forms a multi-protein complex with chaperon proteins such as heat shock proteins (HSP70 and HSP90) and FKBP52 [222]. In the cytoplasm, binding of progesterone to PR stimulates its dissociation from chaperon proteins followed by homo-/hetero-dimerization of PR. These dimers translocate to the nucleus where they bind to the DNA through progesterone receptor element (PRE), AP-1 or SP-1 to induce recruitment of co-regulators (genomic action) [219]. However, PR dimers, present in the cytoplasm or associated with the membrane can activate mitogenic protein kinases such as Src and MAPK (non-genomic action) [218].

Like ER, quantitative assessment of PR expression using IHC staining is routinely performed in all breast cancer patients. [209]. PR, together with ER, is a significant prognostic factor in early breast cancer as it helps assessing the risk of early relapse in patients receiving endocrine therapy [228, 229]. It is also an independent prognostic and predictive marker among ER-positive breast cancer patients [230, 231]. Furthermore, Bardou et al. demonstrated that PR status provides additional value to ER in predicting the benefit of endocrine treatment among patients with primary breast cancer [230]. Another interesting study claims that PR is a stronger predictor than ER regarding the response to adjuvant tamoxifen in premenopausal breast cancer patients [232].

Not all ER-positive breast cancers are found to express PR, as there are ER+/PR- and ER-/PR+ phenotypes. Loss of PR expression often identifies luminal B breast

cancers that are associated with a worse overall prognosis and a higher risk of relapse among ER-positive patients [233]. Also, it was reported that ER+/PR-, ER-/PR+, or ER-/PR- breast tumors have a lower response to tamoxifen and a relatively higher risk of death when compared to ER+/PR+ tumors [234, 235]. This being said, Hefti et al. used gene expression microarray-based analysis of 4000 breast cancer patients to show that ER-/PR+ phenotype is very rare and not a reproducible subtype [236].

5.1.3 HER2 (ERBB2)

HER2 belongs to the HER family of trans-membrane receptor tyrosine kinase that includes the following four members; EGFR (ErbB1, HER1), ErbB2 (HER2), ErbB3 (HER3) and ErbB4 (HER4). The name ERBB is derived from the name of erythroblastic leukemia viral oncogene [237]. All the HER family members share the same structural domains: an extracellular domain, a trans-membrane domain, an intracellular domain and a C-terminal tail [238]. The HER2 gene is located on the long arm of chromosome 17 and has been found to be amplified in several cancers including breast, ovary, endometrium, bladder, lung, colon, and head and neck [51].

The HER2 receptor is present at the cell surface as monomer [239]. Attachment of ligands to HER2 receptor stimulates homo-dimerization with other HER2 receptors or hetero-dimerization with other members of the HER family or other membrane receptors such as insulin-like growth factor receptor 1 (IGF-1) [240]. The dimerization step leads to phosphorylation of tyrosine residues in the cytoplasmic domain, which then activates multiple signaling pathways such as RAS/RAF/MAPK, PI3K/AKT, and PLC/PKC. Activation of these signaling pathways drives a number of different cellular responses such as cell proliferation, differentiation, cell survival and cytoskeletal rearrangement [51, 52].

The last 2014 recommendations of The ASCO/ CAP (College of American Pathologist) stated that HER2 status should be determined in all patients with invasive breast cancer [241]. HER2 status can be assessed using either IHC and/or fluorescent in situ hybridization (FISH). The scoring of HER2 expression levels by IHC is based on both the pattern and intensity of cell membrane staining. IHC scores are used to measure

HER2 expression. Results are reported as 0-1+ (negative), 2+ (equivocal) or 3+ (positive) (Figure 1.7). According to the last recommendations, if the IHC results are equivocal, reflex test should be done using an alternative assay such as FISH. It is also recommended to repeat the test if the results are discordant with other pathologic findings (e.g. 3+ tubular of classical lobular carcinoma) [241].

FISH assay assess the HER2-to-CEP17 ratio and gene copy number (dual-probe ISH). Amplification in a dual-probe ISH assay is defined by examining first the HER2/CEP17 ratio followed by the average HER2 copy number. The interpretation of HER2 FISH scores can be one of the following: positive HER2 amplification when HER2/CEP17 ratio ≥ 2.0 with average HER2 copy number ≥ or < 4.0 signal/cell. Also, HER2 is considered positive when HER2/CEP17 ratio < 2.0 with average HER2 copy number ≥ 6.0 signal/cell. HER2 FISH score is considered to be negative when HER2/CEP17 ratio < 2.0 with average HER2 copy number < 4.0 signal/cell and equivocal when HER2/CEP17 ratio < 2.0 with average HER2 copy number ≥ 4.0 and < 6.0 signal/cell. "Indeterminate for HER2" is a new category of interpretation that indicates the occurrence of technical issues that interfere with accurate interpretation of one or both tests (IHC and FISH). Other in situ hybridization techniques "the bright field ISH" can be used to assess HER2 gene status using a regular light microscope [241]. Of note, HER2 overexpression is not limited to HER2-positive breast cancer as some of luminal breast cancers can also be positive for HER2 [242].

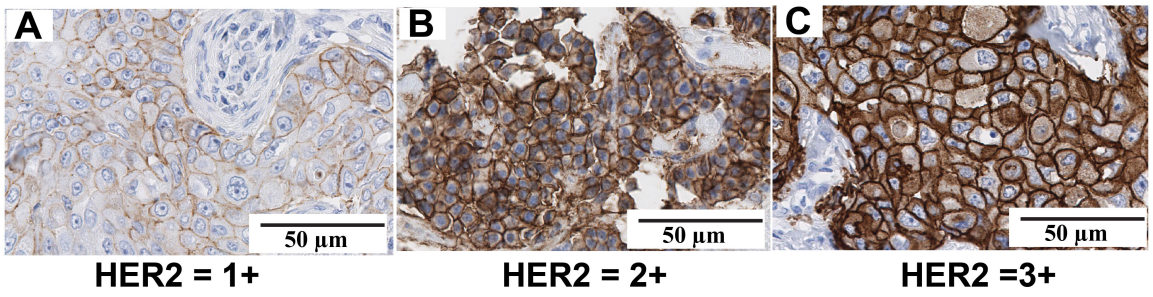


Figure 1.7: IHC scoring of HER2 expression. A) HER2 score = 1+ (negative), B) HER2 score = 2+ (equivocal), C) HER2 score = 3+ (positive)

Amplification of the HER2 gene has been shown to have a great prognostic value in breast cancer. Its overexpression is associated with aggressive breast tumors, a high rate of recurrence, early relapse and increased mortality [243, 244]. The association of HER2 overexpression with shorter disease free and overall survival is independent of other clinical variable such as age, stage and tumor grade [245, 246]. However, HER2 overexpression is significantly correlated with older age > 50 years, high tumor stage and high histological grade of breast cancer [247]. On the other hand, HER2 status is predictive of response to endocrine therapy, adjuvant chemotherapy and HER2-targeted therapy. Enhanced HER2 expression predicts tamoxifen resistance in estrogen receptor positive primary and metastatic breast cancer [248-251]. Furthermore, HER2 status can also predict resistance or sensitivity to different chemotherapeutic agents such as anthracycline, docetaxel and taxane chemotherapy [252-254]. Lastly, overexpression of the HER2 gene identifies primary and metastatic breast cancer patients who will benefit from HER2-targeted therapy such as trastuzumab (Herceptin ®) and tyrosine kinase inhibitors such as lapatinib [57, 58]. Results coming from different studies demonstrated the fundamental role of HER2-targeted therapies in improving clinical outcomes and increasing the number of pathological complete response rates [255, 256].

5.2 Multigene expression assays available in clinical setting

Multigene assays have been introduced with the hope of providing better indicators than conventional clinico-pathologic parameters in terms of classification, prognosis and prediction of breast cancer [99]. In addition, it has been demonstrated that using these assays one could modify the treatment strategy in 25-30% of breast cancer patients [257]. Although these assays use different techniques and different combination of gene sets (Table 1.3), none of them could reach 100% sensitivity and specificity [258].

5.2.1 Oncotype Dx™ (21-gene assay)

Oncotype Dx™ is a reverse transcriptase–polymerase chain reaction (RT-PCR) assay that measures the expression of 21 genes, 16 cancer related genes and 5 reference genes (table 1.4) in RNA extracted from FFPE tissues [166]. In 2004, Paik and coworkers validated the capability of Oncotype DX™ assay to predict the likelihood of distant recurrence in

tamoxifen-treated patients with node-negative and estrogen receptor positive breast cancer (prognostic significance) [259]. Other key studies have shown that Oncotype Dx™ helps predicting the likelihood of 10-year survival (distant relapse free survival DRFS) for breast cancer patients and the magnitude of benefits from adjuvant chemotherapy (predictive significance) [260, 261]. Moreover, promising results have been reported in different studies when the Oncotype DX™ assay has been tested in estrogen receptor positive, lymph node-positive breast cancer patients [262-264]

Oncotype Dx™ calculates recurrence scores (RS) ranging from 0-100 to categorize breast cancer patients (hormone receptor positive, lymph node negative) into three different groups; low-risk (RS < 18), intermediate-risk (RS =18-31) and high-risk (RS ≥ 31) of distant recurrence (Table 1.5). Paik et al. reported distant recurrence rates of 6.8%, 14.3%, and 30.5% in the different risk groups [259]. Likewise, a significant association has been found between RS and the risk for local recurrence in patients with node-negative, ER-positive breast cancer [265]. Furthermore, patients with high RS derive a larger benefit from chemotherapy when compared to those with low RS [261]. However, the extent of benefits in patients with intermediate score remains uncertain [261]. The TAILORx trail (Trial Assigning Individualized Options for Treatment (Rx)) was launched in 2006 to assess the ability of RS to guide therapeutic decisions in breast cancer patients with intermediate score [266].

Oncotype DX™ is performed only in the USA by Genomic Health test center, which is Clinical Laboratory Improvement Amendments (CLIA)-certified and CAP-accredited centralized laboratory [267]. Interestingly, this assay has been incorporated into the ASCO, the National Comprehensive Cancer Network (NCCN) and St Gallen clinical practice guidelines [45, 166, 268]. Overall, oncotype Dx™ assay has many advantages as it enables more individualized treatment decisions in estrogen receptor-positive breast cancer patients, lowers both adverse effects and costs associated with chemotherapy over-treatment and identifies patients who may not have been offered chemotherapy based on traditional clinical and pathological measures alone [269]. However, there are important limitations associated with the use of Oncotype Dx™

including its high cost (4,175.00$ per assay), criteria of inclusion (ER-positive, lymph node-negative) and lack of cross-validation by other independent laboratories [270].

Table 1.3: Multigene expression assays available in clinical settings.

	Oncotype Dx	**MammaPrint**	**PAM50 (Prosigna)**
No. of genes analyzed	21 genes (16 cancer related genes + 5 reference genes)	70 genes	50 genes+ 5 reference genes
Tissue requirement	FFPE tissue	Fresh tissue	FFPE tissue
Technique	qRT-PCR	Microarray-based gene expression profiling	qRT-PCR/Nanostring nCounter
Target patients	ER-positive, lymph node negative patients	Patients with stage 1 or 2 invasive breast cancer, <61 years old, with a node negative tumor less than $5mm^3$	In USA, ER-positive, lymph node-negative, Stage I or II breast cancer OR ER-positive, lymph node-positive (1-3 positive nodes), Stage II breast cancer [271].
Score	• Recurrence score (RS)(0-100) • Predict the risk of 10-years distance recurrence Low-risk (RS < 18), Intermediate-risk (RS =18-31), High-risk (RS ≥ 31)	Predict the risk of 10-year distant metastasis-free survival rates Low risk (good prognosis), High risk (bad prognosis)	• Risk of recurrence (ROR) • Predict late distant recurrence in the first 5 years Low risk (ROR < 29), Moderate risk (ROR = 29-53), High risk (ROR ≥ 53) • Identify different molecular subtypes of breast cancer
Performance site	Central Genomic Health test center, USA	Central Agendia, Amsterdam, the Netherlands	Not central The most known is ARUP Laboratories
Guideline	NCCN/ASCO, St Gallen clinical practice guidelines	St Gallen clinical practice guidelines, FDA cleared	Received the CE mark, FDA cleared

Table 1.4: Panel of genes in Oncotype Dx™ assay (21-gene assay) [166].

Groups of genes	Gene name	Gene description
Estrogen-related genes	ER PR Bcl2 SCUBE2	Estrogen receptor Progesterone receptor B-Cell CLL/Lymphoma 2 Signal Peptide, CUB and epidermal growth factor-Like Domain-Containing Protein 2
HER2-related genes	HER2 GRB7	Human epidermal growth factor receptor 2 Growth factor receptor-bound protein 7
Proliferation genes	MKI67 STK15 (Aurora A kinase) BIRC5 (Survivin) CCNB1 (Cyclin B1) MYBL2	Marker of proliferation Ki-67 STK15: Serine/Threonine-Protein Kinase 15 Baculoviral inhibitor of apoptosis repeat-containing 5 G2/mitotic-specific cyclin-B1 V-Myb Avian Myeloblastosis Viral Oncogene Homolog-Like 2
Invasion genes	MMP11 (Stromolysin) CTSL2	Matrix metalloproteinase 11 Cathepsin L2
Macrophage genes	CD68	Cluster of differentiation 68
Others	GSTM1 BAG1	Glutathione S-transferase Mu 1 BCL2-associated athanogene
Reference genes	ACTB GAPDH RPLPO GUS TFRC	β-actin Glyceraldehyde 3-phosphate dehydrogenase Large ribosomal protein β-glucuronidase Transferrin receptor protein 1

Table 1.5: Oncotype Dx™ recurrence score

Recurrence score (RS)	Risk of recurrence	Benefit from chemotherapy	Likelihood of 10-year DRFS
<18	Low	Minimal or no benefit	High
18-30	Intermediate	Uncertain	Intermediate
≥30	High	High	Low

DRFS: distant relapse free survival

5.2.2 MammaPrint (70-gene assay)

MammaPrint is a microarray-based gene expression assay of RNA extracted from fresh breast cancer tissue performed by Agendia, which is a central laboratory in Amsterdam, the Netherlands [166]. This assay studies the expression of 70 genes that are associated with different hallmarks of cancer such as genes regulating cell cycle, cell growth, proliferation, invasion and metastasis, adaptation to unfamiliar microenvironment as well as angiogenesis (Appendix I) [156, 272]. MammaPrint is currently recommended for patients with stage 1 or 2 invasive breast cancer, < 61 years old, with a node negative tumor less than $5mm^3$ [270]. It is a prognostic tool that categorizes early-stage breast cancer patients as low risk (good prognosis) or high risk (bad prognosis) groups that correspond to 10-year distant metastasis-free survival rates of > 90% or < 90%, respectively [273]. Only patients with high risk of metastasis need supplementary chemotherapy in contrast to those with low risk of recurrence who can forego adjuvant chemotherapy with its associated toxicity (Predictive significance) [274].

Microarray In Node-Negative and 1 to 3 positive lymph node Disease may Avoid Chemotherapy (MINDACT) trail evaluates the ability of MammaPrint to select the right breast cancer patients for treatment with chemotherapy in addition to surgery and hormonal therapy. In 2007, MammaPrint has received clearance for use by the FDA. In addition, this assay is included in the clinical practice guidelines of St. Gallen for assessing risk of distant relapse and the benefit of adjuvant chemotherapy [45]. However, both the ASCO the European Society for Medical Oncology (ESMO) reported that the final results from the MINDACT trial, which will not be available for several years, are crucial to prove the MammaPrint as an important prognostic and predictive tool in breast cancer treatment [166, 275]. As for Oncotype Dx™, MammaPrint is costly (4,200$ in the USA). In addition, the requirement of relatively large amount of fresh tissue that has to be snap-frozen in liquid nitrogen within one hour after surgery is another challenging point for MammaPrint assay [166].

5.2.3 PAM50 (50-gene assay, Prediction Analysis of Microarray 50)

The PAM50 Breast Cancer Intrinsic Classifier™ is an RT-PCR assay that studies the expression of 50 classifier genes and five control genes to identify different molecular subtypes of breast cancer; luminal A, luminal B, HER2-enriched and basal-like (Appendix II) [276]. The normal-like molecular subtype is not reported in the commercial PAM50 assay as research microarray studies reported that this subtype is due to cancer specimens containing too much normal breast tissue [157, 277]. Furthermore, PAM50 calculates the risk of recurrence (ROR) by weighing the molecular subtype in combination with clinical variables such as tumor size, grade, histological subtypes, ER status and node status. Patients were categorized based on ROR into three groups; low risk (ROR < 29), moderate risk (ROR = 29-53) and high risk (ROR ≥ 53) [40]. This assay uses RNA extracted from FFPE tissues.

PAM50 has been extensively studied in different research settings. Kelly and his colleagues have evaluated the performance of Oncotype DX and PAM50 among patients with early-stage ER-positive breast cancer. They reported a good agreement between the Oncotype DX and PAM50 assays for high (i.e., luminal B and RS ≥ 31) and low (i.e., luminal A and RS <18) prognostic risk assignment. In addition, about half of the intermediate RS group was reclassified as low risk by the PAM50 that could forgo some patients from receiving chemotherapy [277]. Other studies confirmed that PAM50 assay gives more prognostic information than standard methods such as clinical factors (tumor size, nodal status, histologic grade, patient's age) and IHC study of ER, PR, HER2 and Ki-67 biomarkers in tamoxifen-treated estrogen receptor-positive breast cancer [157].

PAM50 assay is highly predictive of neoadjuvant response when considering all breast cancer patients including those with ER-negative disease [40]. Prat et al. demonstrated that PAM50-ROR was consistently found to be independent predictors of relapse and its combination with OncotypeDX, Mammaprint and SET (index of sensitivity to endocrine therapy) will significantly increase the performance of prediction [278]. Interestingly, the PAM50 based-Prosigna breast cancer prognostic gene signature assay received the CE mark in 2012, which indicates European approval followed by the U.S. FDA clearance in 2013 [279].

5.3 Proliferative biomarkers in breast cancer

Uncontrolled cell proliferation is one of the hallmarks of cancer [6]. There is now growing evidence that the degree of proliferative capacity of breast cancer impacts on its aggressiveness, prognosis and clinical behavior [280-282]. Various approaches have been developed to assess proliferation in breast cancer such as DNA microarray, mitotic score, and immunohistochemistry of proliferation-associated antigen such as Ki-67, cyclin A, MCMs and topoisomerase IIα [112, 281, 283]. However, to date, mitotic score remains the only proliferative marker incorporated in routine clinical practice [284]. In the following section, we will focus on mitotic score and two IHC proliferative markers Ki-67 and MCM2.

5.3.1 Mitotic score

Mitotic count is one of the three components of breast tumors grading system [285]. Counting mitotic figures is the simplest and the most conventional approach in assessing proliferative capacity of breast cancer cells [286]. Mitotic count can be achieved by calculating the number of mitotic figures in 10 consecutive high power fields (x400) in routine H&E stained slides [287]. Mitotic counting has been described as a powerful, easy to perform, cheap, reliable and reproducible proliferation marker that can be applied to all cases of breast cancer [288, 289]. In 2013, the 13th St Gallen International Breast Cancer Conference Expert Panel included histological grading as one of the factors to determine inclusion of chemotherapy in treatment of luminal A breast cancer [45].

Different studies have reported the importance of the mitotic count in weighing the prognosis of breast cancers [287, 290-292]. It has been reported that a low mitotic index in the primary breast tumor is associated with better prognosis and less aggressive biological behavior than tumors with high mitotic index [293, 294]. Moreover, a high mitotic count in the primary breast tumors is correlated to a higher incidence to develop metastases [295]. Evaluating proliferative activity by counting mitotic figures in lymph node metastases of breast cancer patients also provides some prognostic values as it can predict the clinical course of distant metastasis and the clinical outcome of those patients [296-298]. In addition, Penault-llorca et al. demonstrated variation in the mitotic index

before and after neoadjuvant chemotherapy that bears on the prognosis [299]. Besides, high mitotic index is predictive to pathologic complete response to neoadjuvant chemotherapy, meaning that patients whose tumors have high mitotic count respond better to neoadjuvant treatment than those patients whose tumor have low mitotic count [300, 301].

5.3.2 Ki-67

Ki-67, a nuclear non-histone protein was originally described by Gerdes and his colleagues in the 1980s. It was characterized using a monoclonal antibody directed against a Hodgkin's lymphoma derived cell line. This antibody was coined after the city of Kiel in Germany where the clone was raised while the suffix 67 refers to the clone number on the 96-well plate [302, 303]. The expression of Ki-67 in all proliferating cells and its absence in quiescent cells raised great interest on its potential role as a marker of cell proliferation [304]. However, afterwards other studies demonstrated that proliferating cells that pass from G0 to G1 phase of the cell cycle do not express Ki-67 and that the onset of Ki-67 expression in those cells occurs only in the late G1 phase [305, 306]. In addition, there is alteration of Ki-67 concentration during the different phases of the cell cycle with the highest concentration found during the G2 and S phases and the nadir during anaphase and telophase of the cell cycle. Low levels of Ki-67 could also be detected in G1 phase of the cell cycle [307-309].

Taken together, it seems that Ki-67 plays a crucial role during cell division but its exact function inside the cell is currently unknown [281]. Notably, despite the fact that Ki-67 depleted cells and their cognate nuclei are smaller than their normal counterpart, these cells maintain a normal cell-cycle profile [310]. In addition, there is evidence that Ki-67 might be involved in other non cell cycle related processes such as ribosomal biosynthesis [311, 312]. Trihia et al. demonstrated that Ki-67 allows estimation of the growth fraction of breast cancer cells [286]. Furthermore, it has been reported that Ki-67 is more accurate than mitotic score in assessing tumor cell proliferation and predicting prognosis of breast cancer patients [313].

Many studies have looked into the prognostic significance of Ki-67 expression in breast carcinomas [286, 314-316]. High Ki-67 expression in breast cancer is associated

with ER negativity, HER2 positivity and high histological grade [317]. It has been reported that elevated expression levels of Ki-67 in early breast cancer is a strong predictor of a higher risk of relapse and reduced disease-free survival [314, 318]. Moreover, Ki-67 overexpression is more likely to be associated with axillary nodal metastasis in high-grade breast carcinoma [319]. Interestingly, most axillary lymph node metastases display higher levels of Ki-67 expression than that of the primary breast tumor suggesting that metastatic cells have a higher proliferating potential and more aggressive behavior than the primary breast cancer cells [320].

On the other hand, Ki-67 is a predictive marker of response to endocrine therapy as high levels of Ki-67 expression identify patients that particularly benefits from hormonal therapy [316]. In addition, high Ki-67 index is a predictive factor of pathologic complete response to neoadjuvant chemotherapy in patients with locally advanced breast cancers [299, 321].

Another important issue that has been thoroughly studied is the prognostic value of endocrine and chemotherapy induced variation in Ki-67 expression level. Breast tumors with decreased Ki-67 expression following endocrine or chemotherapy were more likely to achieve complete pathological response [322, 323]. Ki-67 index of $\geq$ 15% after preoperative chemotherapy is a significant poor prognostic indictor in breast cancer patients as it is associated with a worse disease-free survival and a higher incidence of recurrence compared to other patients [324]. Data from the IMPAKT study have shown that elevated levels of Ki-67 two weeks after the onset of endocrine therapy was predictive of a worse recurrence-free survival [325]. Dowsett and his colleagues demonstrated that combining Ki-67 measurements in breast cancer patients before and after endocrine therapy could improve the prediction of recurrence-free survival [326].

Luminal breast cancer (ER+/HER2-) can be classified into luminal A and luminal B breast cancer based on 14% Ki-67 index [50]. However, the IMPAKT 2012 working group stated that there was not enough evidence to support a Ki-67 labeling index of 14% to identify clinically relevant subtypes of breast cancer [112]. Furthermore, further studies suggested different Ki-67 cutoff points such as 10%, 14% and 20% to distinguish luminal A from luminal B breast cancer [45, 50, 282, 327-329]. In 2013, the 13th St Gallen International Breast Cancer Conference Expert Panel adopted a new approach to

the classification of breast cancer patients for therapeutic purposes and recommended to use a 20% Ki-67 index to discriminate between luminal A and luminal B breast cancers [45].

So far, Ki-67 has failed to impose itself as an accepted breast cancer biomarker because of marked heterogeneity, poorly standardized staining and scoring conditions, lack of reproducibility and the difficulty to establish an appropriate cutoff point [45, 50, 282, 327-329]. Due to the above limitations, neither the ASCO nor The NCCN recommended to use Ki-67 as a routine breast cancer biomarker [281, 330]. Surely, there is much more research to be done to propose the use of Ki-67 as a standard proliferative marker in clinical laboratories.

5.3.3 MCM2

MCM2 is one of the highly conserved family members of minichromosome maintenance (MCMs) proteins that function together as a complex in the cell [331]. MCM2-7 heterohexameric ring complex plays a crucial role in initiation of DNA replication and unwinding of the DNA due to its replicative helicase activity [332, 333]. It is one of the vital mechanisms that ensure single replication of the DNA per cell cycle [334]. During G1 phase of the cell cycle, pre-replication complex (pre-RC) is assembled by interaction of the following proteins at the replication origins: origin recognition complex (ORC), CDC6, CDT1 and MCM heterohexameric ring complex. During S phase, MCM complex unwind the double stranded DNA to allow initiation of replication then move out of the replicated DNA while remaining attached to un-replicated DNA. MCMs dissociate from the ORC and remain in the nuclear matrix during G2 and early mitosis until a new round of replication starts again (Figure 1.8). If the cells decide to exit the cell cycle and enter G0, MCMs are displaced from chromatin and degraded, as this complex is no longer required [311, 334]. Of note, there is a growing body of evidence to suggest that MCM proteins exert important roles in chromosomal dynamics and integrity, genomic stability and DNA repair [335, 336].

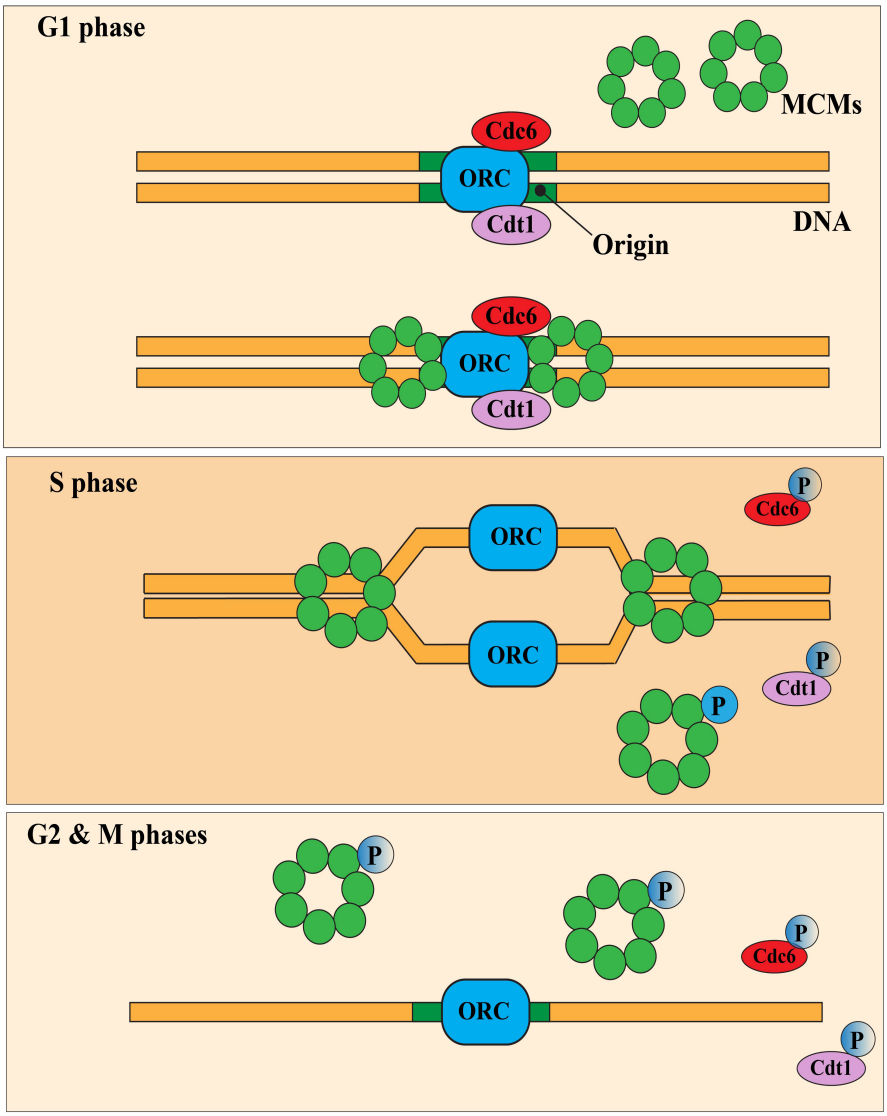


Figure 1.8: Role of MCMs complex in precise DNA replication

During G1 phase of the cell cycle, pre-replication complex (pre-RC) is assembled by interaction of the following proteins at the replication origins: origin recognition complex (ORC), CDC6, CDT1 and MCM heterohexameric ring complex. During S phase, MCM complex unwind the double stranded DNA to allow initiation of replication then move out of the replicated DNA while remaining attached to un-replicated DNA. MCMs dissociate from the ORC and remain in the nuclear matrix during G2 and early mitosis until a new round of replication starts again [311, 334]

Considering the above-mentioned functions of MCMs, MCM2 has been considered a potential prognostic biomarker in different types of cancer such as colon, prostate, lung, lymphoma and breast [337-340]. Overexpression of MCM2 is significantly associated with poor outcomes in diffuse large B cell lymphoma and in non-small cell lung cancer [340, 341]. In addition, MCM2 is an independent predictor of recurrence in prostate cancer and hepatocellular carcinoma [342]. It is also one of the candidate genes identified by Cunha et al. to control local aggressiveness and metastatic behavior in soft tissue tumors [343]. Furthermore, Gonzalez and his colleagues showed that MCM2 is a strong independent prognostic marker in breast cancer as it is significantly associated with poor survival and development of regional and distant metastases. These authors indicated that high expression of MCM2 is associated with tumor size, higher mitotic index and a higher histologic grade of breast cancer [344].

MCM2 can also be predictive; for instance it may help to determine which cervical tumors are likely to respond to radiation therapy [345]. Loddo et al. provided evidence that MCM2 has a predictive value in breast cancer, as tumors with elevated levels of MCM2 are more likely to benefit from chemotherapy [346]. The fact that MCM2 is present in all proliferating cells and at all stages of the cell cycle together with its absence in non-proliferating cells support the notion that MCM2 is an ideal proliferative biomarker [347]. Many studies have reported its role as proliferative marker in numerous malignant tumors such as prostate, thyroid and colorectal carcinoma [348-350]. In addition, there was a positive correlation between MCM2 and other proliferative markers such as Ki-67 and PCNA in colorectal carcinoma and in ductal breast carcinoma [349, 351]. Finally, MCMs together with other members of Pre-RC could be attractive targets to design new cancer therapies [311].

5.4 Other emerging biomarkers in breast cancer

5.4.1 Annexin A1 (ANXA1)

Annexin A1, is a 37KDa protein that belongs to the superfamily of calcium-dependent phospholipid-binding proteins that share a peculiar protein core domain [352]. Human annexin A1 is translated from the ANXA1 gene, which is located on chromosome

9q21 [353]. ANXA1 is known to be present in the cytoplasm, membrane and nucleus. However, its subcellular localization is critically dependent on certain stimuli; for example, ANXA1 translocate to the membrane and exported to the extracellular surface of the cell membrane in response to glucocorticoids [354]. In contrast, epidermal growth factor (EGF) stimulation and stress induce its translocation to the nuclei [355, 356]. Interestingly, nuclear localization of ANXA1 has been reported to correlate with advanced gastric carcinoma and with peritoneal dissemination [357].

ANXA1 is thought to function as an anti-inflammatory molecule [358]. However, there is growing evidence that it participates to other critical cellular processes implicated in development and progress of cancers such as cell proliferation, differentiation and apoptosis [359-361]. It has been reported that ANXA1 regulates the epithelial cell migration in some malignancy [362]. Furthermore, loss of ANXA1 has been associated with an increased susceptibility to DNA damage and mutations [363]. Several studies have confirmed a deregulated ANXA1 expression in a variety of tumors. Overexpression of ANXA1 has been reported in gastric, hepatic and pancreatic carcinomas [364-366]. On the other hand, decreased expression of ANXA1 has been described in other types of cancers as in prostate and thyroid carcinomas [367, 368].

More specifically, differential expression of ANXA1 in normal and malignant mammary cells has been reported in different studies. However, the results of these studies are somewhat controversial and confusing with regards to two important issues. The first one is the level of ANXA1 expression in breast cancer tissue when compared to normal breast tissue as some of these studies reported increased expression of ANXA1 in breast cancer [369, 370], while others showed the opposite [371-373]. The second issue is the role of ANXA1 in progression and metastasis of breast cancer cells. Several studies have claimed that ANXA1 increases the metastatic potential of tumors [374, 375]. In contrast, others studies reported that ANXA1 attenuates epithelial-mesenchymal transition and reduces metastatic potential in breast cancer [376].

ANXA1 has also been demonstrated as a potential prognostic biomarker in different types of cancer such as lung, breast, gastric, colon and nasopharyngeal carcinoma [366, 377-379]. Increased expression of serum ANXA1 was significantly correlated with pathological grade and clinical stage of lung cancer patients [377].

Moreover, up-regulated ANXA1 expression is significantly associated with invasion, lymph node metastasis and poor prognosis of gastric and colon cancer patients [366]. Also, elevated levels of ANXA1 has been reported to be a promising marker to predict the recurrence of bladder cancer [378]. In addition, high expression of ANXA1 was associated with unfavorable disease characteristics in invasive breast carcinoma such as high histological grade [375, 379]. Lastly, a panel of four genes that includes ANXA1, PRKCA, DUSP2 and SERPINA3 has recently been suggested to be a reliable predictive tool for the identification of breast cancer response to neoadjuvant chemotherapy [380].

5.4.2 MMP-9

Matrix metalloproteinase-9 (MMP-9) is a zinc-dependent peptidase that belongs to the gelatinase subfamily of MMPs. It is excreted as an inactive pro-enzyme that undergoes activation upon cleavage by different types of extracellular proteases [381]. MMP-9 is regulated through interactions with growth factors and cytokines and activation of intracellular signaling pathways [382]. In transgenic mice, MMP-9 expression is elevated in osteoclasts and migrating keratinocytes [383]. Interestingly, in many normal human tissues and human tissue cell lines (epithelial, connective tissue, and muscle origin) MMP-9 in not expressed [384, 385]. According to several studies, MMP-9 is produced in many stromal cells such as fibroblasts, macrophages, granulocytes and lymphocytes adjacent to the tumor cells [386].

MMP-9 acts mainly to degrade gelatin present in the ECM and collagen type IV of the basement membrane implying an important physiological role in tissue remodeling embryonic implantation and wound healing [387, 388]. In addition, MMP-9 allows implantation primary tumors and contributes to early stages of tumor growth, cell proliferation and angiogenesis [389-391]. MMP-9 also contributes to cancer dissemination and invasiveness as it affects tumor cellular adhesion, remodeling of the ECM and cleavage of membrane-bound growth factors. Because MMP-9 allows the establishment of pre-metastatic niches, it paves the way to tumor cell dissemination and facilitates the growth of metastatic cells in the distant organs [387, 391-394].

MMP-9 is also a promising prognostic biomarker in various types of cancer including cervical, colorectal and ovarian neoplasms for elevated serum and urine levels

of MMP-9 are associated with invasion, metastasis and a poor prognosis [395-400]. In addition, MMP-9 has been described as a poor prognostic factor in breast cancer patients being associated with higher TNM stage, lymph node involvement and a high propensity to distant metastasis [401, 402]. Elevated levels of MMP-9 are correlated with triple negative and HER2-positive subtypes of breast cancers that are known to have high metastatic potentialities [403, 404]. Moreover, high level of MMP-9 is significantly correlated with shorter time to relapse and a lower survival rate in breast cancer patients [405]. Of note, MMP-9 can also be envisaged as a potential predictive marker for adjuvant systemic therapy in breast cancer patients [406]. Lastly, it is interesting to note that elevated serum level of MMP-9 before the initiation of neoadjuvant chemotherapy for locally advanced breast cancer has been correlated with improved clinical response [407].

6. Objectives of this work

There are many evidences that each breast cancer has its own set of genomic abnormalities or altered pattern of gene expression that can explain the aggressiveness of each tumor, its ability to metastasize and its response to chemotherapeutic agents or other forms of targeted therapies. In this study, our general objective is to identify and characterize new biomarkers with prognostic value in aggressive subsets of breast cancer focusing primarily on triple-negative tumors and luminal B breast cancer. In this thesis, we studied two biomarkers, ANXA1 and MMP-9 as potential prognostic biomarkers for triple-negative breast cancer and MCM2 as a potential proliferative biomarker that can distinguish between luminal A and luminal B breast cancer.

I. Novel biomarkers in triple-negative molecular subtypes of breast cancer

A) ANXA1: our aim was to investigate the differential expression of ANXA1 among various histological grades and molecular subtypes of breast cancers and in particular triple negative tumors. Also, we sought to study the putative role of ANXA1 in EMT by comparing its expression with a panel of EMT markers.

B) MMP-9: The aim of this study was to assess the potential clinical usefulness of MMP-9 as a prognostic biomarker of breast cancer. We also wanted to study the expression of MMP-9 in different histological grades and molecular subtypes of breast cancer.

II. Novel biomarkers to distinguish luminal A from luminal B molecular subtypes of breast cancer

MCM2: The main aim of this study is to determine if MCM2 can be an alternative to Ki-67 for measuring breast cancer cell proliferation. We studied the expression of MCM2 in comparison to Ki-67 as they relate to breast cancers of different histological grades and molecular subtypes focusing primarily on ER-positive tumors.

CHAPTER II

Deregulated Expression of ANXA1 in Human High-Grade Breast Cancers

Einas M Yousef[1], David Laperriere[1], Muhammad R Tahir[2], Sylvie Mader[1&3], Louis A Gaboury[1&4§]

[1]Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, Canada, [2]Le centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Canada, [3]Department of Biochemistry, Faculty of Medicine, Université de Montréal, Montréal, Canada, [4]Department of Pathology and Cell Biology, Faculty of Medicine, Université de Montréal, Montréal, Canada

§Corresponding author

Published in Journal of molecular biomarkers and diagnosis, November 2013

Authors' contributions

EY and LG are responsible for the study design.

EY performed the experiments and collected the data.

DL performed the *in silico* analysis.

EY and LG scored the IHC reactions.

EY prepared all the IHC diagrams and pictures, DL prepared the *in silico* analysis diagrams.

MT carried out the statistical analysis.

EY, SM and LG participated in the data analysis and interpretation.

All authors drafted, read and approved the final manuscript.

Abstract

Context: Breast cancer is a heterogeneous disease comprising a diversity of tumor subtypes. All of them differ from each other resulting in a wide array of clinical pictures, risks of recurrence and response to treatment. In recent years, tumor biomarkers have changed the way breast cancers are diagnosed and treated.

Objectives: In this paper, we have sought to investigate the differential expression of ANXA1, a multifunctional calcium-binding protein, among various molecular subtypes of breast cancers and in particular triple-negative tumors.

Design: ANXA1 was first studied using *in silico* analysis on available DNA microarray and RNA sequencing data of human breast tissues. Next we ascertained ANXA1 expression on cell lines and breast carcinoma tissue microarrays along with cognate normal breast tissue.

Results: Whereas ANXA1 expression is normally restricted to the normal myoepithelial cell layer it becomes ectopically and aberrantly expressed in tumor cells of a significant minority of aggressive breast cancers. Specifically, we found that ANXA1 expression is severely deregulated in high-grade breast cancers that comprise clinically aggressive tumors such as triple-negative and, to some extent, HER2-positive breast cancers.

Conclusion: Our results indicate that ANXA1 is a valuable breast cancer biomarker that can help to segregate and dissect out subsets of high histological grade breast cancers paving the way to a better understanding of breast cancer progression and metastasis.

Keywords: ANXA1; Multifunctional calcium binding protein; Breast cancer

Introduction

Breast cancer is a heterogeneous disease characterized by a wide range of pathological features, diverse clinical behavior and variable response to treatment based on a complex interplay between characteristics of the tumor and the host [23]. In 2013, approximately 1 in 8 (12%) women in the USA will develop invasive breast cancer. Breast cancer is the second cause of death after lung cancer among American women. The American Cancer Society's estimated 39,620 deaths from breast cancer patients to occur this year [408]. The high mortality rate from breast cancer results mainly from metastasis due to spread of cancer cells to distant organs such as the liver, lungs, brain and bones [22]. In selective subtypes of breast cancer, therapies targeting specific signaling pathways are well known and many of them are now widely used in the clinics [409]. For instance, hormonal therapy (Tamoxifen and aromatase inhibitors) can inhibit the effect of estrogen or decrease estrogen levels in patients with hormonally responsive tumors. In HER2-positive patients, Trastuzumab can prolong survival and lower the risk of relapse. Unfortunately, patients whose tumor fail to express hormonal receptors or lack HER2 overexpression will not benefit from those types of treatment and hence must rely mainly on chemotherapy. There is now a large body of evidence to suggest that cancer treatment must be tailored to individual tumor characteristics [410]. Despite the success of current therapies, we still need to uncover unique genetic alterations or tumor characteristics that might be translated into prognostic & predictive biomarkers or pharmacologically amenable targets.

Annexin A1 (ANXA1), the first characterized member of the annexin superfamily, is a calcium and phosopholipid binding protein known to mediate the anti-inflammatory actions of glucocorticoids [352]. However, it possesses many other functions that impact on key cellular processes such as proliferation, differentiation, cytoskeletal organization, cell migration and apoptosis [359, 360, 362, 375]. Expression of ANXA1 has been associated with the development and progression of different types of cancer. On the one hand, ANXA1 has been shown to be down-regulated in esophageal cancer, head and neck cancer, and prostate cancer [367, 411, 412]. On the other, ANXA1

was found to be up-regulated in other types of malignancies such as pancreatic cancer, hepatocellular carcinoma and stomach cancer [364, 365, 413].

We became interested in the status of ANXA1 gene expression in breast carcinomas following the seminal work by Perou et al. on the molecular classes of breast cancer [36]. Although the authors did not expand specifically on the putative role of ANXA1 in breast cancer, molecular profiling of tumors clearly indicated that ANXA1 was related to the basal-like subtype of breast cancer. Upon carefully reviewing the literature we found that the status of ANXA1 expression and its role in initiation and progression of breast cancer was still an unresolved issue. It was reported that a lower expression of ANXA1 was significantly associated with advanced stage of breast cancer and poor overall survival when compared to patients with high ANXA1 expression [371]. Likewise, it has also been shown that expression of ANXA1 is down-regulated in metastatic tumors suggesting that ANXA1 functions as an epithelial-mesenchymal transition (EMT) /metastatic suppressor [376]. These observations are hard to reconcile though with the findings by Yi & Schnitzer [414] who reported that ANXA1 null mice had impaired tumor growth and a lower propensity to develop metastasis. Furthermore, it was reported by others that ANXA1 promotes metastasis formation of basal-like breast cancer through its regulation of TGF-β [375]. Indeed, Yom et al. demonstrated that high ANXA1 expression was significantly correlated with unfavorable prognostic factors such as hormone receptor negativity, HER2 overexpression and triple-negative breast cancer [379].

In this study, we investigated ANXA1 expression using *in silico* analysis on available DNA microarray and RNA sequencing data of both normal human breast tissue and in a variety of human breast cancers. We also carried out immunohistochemical analyses on a large number of normal breast tissue and breast carcinoma using cell microarrays (CMA) and tissue microarrays (TMA). This allowed us to compare ANXA1 expression at the protein level with that of other available breast biomarkers. Here, we show that ANXA1 is differentially expressed in normal human breast tissue and in breast tumors. Moreover, we found that ANXA1 expression is severely deregulated in high-grade breast cancers that comprise clinically aggressive tumors such as triple-negative and to some extent HER2-positive breast cancers. Our results suggest that ANXA1 is a

valuable breast cancer biomarker whose deregulated pattern of expression may help to segregate subsets of high histological grade breast cancers such as triple-negative category and HER2-positive breast tumors into clinically meaningful categories.

Material and methods

In silico Analysis

The web application bc-GenExMiner [415] was used for correlation analysis of ANXA1 gene expression on a dataset comprising over 3400 microarrays. The "aov" and "TukeyHSD" functions were carried out to compare the mRNA levels within each breast cancer molecular subtypes for the ANOVA and Tukey multiple comparisons of means. The ANOVA was applied to check for an overall difference of expression level between molecular subtypes. The Tukey multiple comparisons of means were used to test for a significant difference between two subtypes (e.g. luminal A vs. basal). For both tests, a *p*-value < 0.05 was considered significant. The mRNA levels of ANXA1, Vimentin (VIM), ESR1 and FOXA1 were compared using publically available microarrays and mRNA sequencing breast cancer patients datasets. The former included 51 breast cancer cell lines [416] and 247 patients [417]. The mRNA sequencing dataset from the Cancer Genome Atlas Network [43] included 844 patients. Moreover, MiSTIC dataset was also used to correlate ANXA1 with a different set of genes in normal breast tissue (102 cases) and breast cancer (756 patients) based on RNA-sequencing data derived from The Cancer Genome Atlas (TCGA). Pearson correlation coefficient between set of genes was calculated with the "cor.test" function of the R language and environment for statistical computing (http://www.R-project.org/).

Cell Culture

All cell lines were purchased from the American Type Culture Collection (ATCC). MCF10F and MDA-MB-231 cells were maintained at 37°C in Dulbecco's modified Eagle's medium Nutrient Mixture F-12 (DMEM/F12) (Wisent) supplemented with 10% fetal bovine serum (FBS) (Sigma), 0.5 µg/ml hydrocortisone, 10 ng/ml EGF, 10 µg/ml insulin, 1% penicillin-streptomycin and 100 ng/ml choleratoxin (Wisent). 184-

B5 cell line was maintained in the same medium except that 1 ng/ml of choleratoxin was used instead of 100 ng/ml. MCF-7 cell line was maintained in Dulbecco's modified Eagle's medium Nutrient Mixture (DMEM) supplemented with 10% FBS and 1% penicillin-streptomycin. ZR-75-1 cell line was maintained in RPMI with 10% FBS and 1% penicillin-streptomycin. MDA-MB-361 was maintained in L15 + 2mM Glutamine + 15% FBS. MDA-MB-468 was maintained in Ham's-F12 medium (1:1 mixture) supplemented with 2 mM L-glutamine and 5% FBS.

Preparation of Paraffin-Embedded Cell Pellets

All cell lines were maintained in culture until they reached 80% confluence. Petri dishes were then washed with PBS followed by trypsinization of the cell layer until they completely detached from the bottom of the dish, then media specific for each cell line was added to inactivate trypsin. Cells and media were collected in 15 ml tube and centrifuged at 1000 rpm for 5 minutes. The supernatant was aspirated and the pellets were resuspended in 10% neutral buffered formalin and kept at 4 °C for 24 hours to allow proper fixation. After 24 hours, the cells were centrifuged at 1,800 rpm for 10 minutes and the supernatant was aspirated and the pellets resuspended in melted HistoGel (thermoscientific). Cell pellets were maintained in HistoGel at 4 °C for 20 minutes until the gel solidified. Pellets were then removed from the eppendorf tubes and placed in embedding cassettes for further fixation in 10% neutral buffered formalin. Cell pellets were next processed in SAKURA Tissue-Tek VIP (Vacuum infiltration processor) then embedded in paraffin blocks. 3 μm section from each block was prepared and stained with Haematoxylin and Eosin (H&E) to evaluate the relative position of each cell cluster in each block and to determine cellularity.

Cell Microarray (CMA) Construction

CMA was prepared as follows: two punches (2mm) from each cell line were plucked and inserted into recipient blocks according to the intended design of the map using a Manual Tissue Arrayer I (Beecher Instruments). Blocks were next inverted and incubated overnight in the oven over a glass slide. The blocks were allowed to cool off until they could be detached from the glass slide. 3 μm sections were prepared from the

CMA and stained with H&E or submitted to automated immunohistochemical (IHC) reactions.

Patient Samples and Tissue Microarray (TMA) Construction

This study was conducted on 300 archived formalin-fixed, paraffin-embedded (FFPE) samples containing both normal and tumor tissues obtained from female patients. All samples were obtained from *Centre Hospitalier de l'Université de Montréal* (CHUM) after obtaining the approval of the research ethical committee (SL 05.019). The collected blocks contained both invasive and *in situ* carcinoma of the breast obtained at surgery (lumpectomy or mastectomy). Normal breast tissues were obtained from healthy women undergoing plastic surgery (Table 2.1). A number of extraneous tissues such as colon, thyroid and placenta were included in each TMA to serve as external controls.

To construct TMAs, 4 μm sections from each paraffin donor block were stained with H&E. These sections were microscopically examined by two independent pathologists to select the most representative fields. Areas from each corresponding paraffin blocks were plucked in duplicate or triplicate and 1 mm core punches realigned into recipient blocks using a Manual Tissue Arrayer I (Beecher Instruments). Tissue sections from each TMA were prepared and one slide from each block was stained with H&E to review the diagnoses and histological grades on all tissue samples.

Histological Grading of Breast Cancer

Modified Scarff-Bloom-Richardson-Ellis-Elston grading system (SBR-EE) [418] was used to score invasive breast cancers. The patients were classified into one of three different histological grades (Grade I, II & III). Briefly, tumor grade is determined based on three parameters: - extent of tubule formation, mitotic rate and nuclear size. Each of these features is assigned a score ranging from 1-3. Next, the scores for each criterion are added together with a cumulative figure ranging from 3 to 9. Grade I tumors are well differentiated (low grade) with a total score of 3 to 5. Grade II tumors are moderately differentiated (intermediate grade) with a total score of 6 to 7. Grade III are poorly differentiated (high grade) with a total score of 8 to 9.

Table 2.1: Clinico-pathological data of tissues used in TMAs

Variables	No. of cores	%
Organs used in TMA	643	
Colon	4	0.7
Lymph node	26	4.0
Mammary gland	596	92.7
Rectum	11	1.7
Thyroid	6	0.9
Grade	570	
I	54	9.5
II	116	20.3
III	400	70.2
Molecular subtypes	488	
Luminal A	143	29.3
Luminal B	54	11.1
HER2-positive	68	13.9
Triple-negative	223	45.7

Immunohistochemistry

Immunohistochemistry was carried out according to manufacturer recommendations on an automated immunostainer (Discovery XT system, Ventana Medical Systems, Tucson, AZ). Antigen retrieval was performed with proprietary reagents followed by incubation with the primary antibody. Antibodies used are listed in table (2.2). Sections were then incubated with a specific secondary biotinylated antibody for 32 minutes. Streptavidin horseradish peroxidase, and 3,3-diaminobenzidine were used according to the manufacturer's instructions (DABmap detection kit, Ventana Medical Systems). Finally, sections were counterstained with Gill hematoxylin and sodium bicarbonate. For ANXA1, MDA-MB-231 cell line and thyroid carcinomas were used as positive controls while MCF-7 and ZR-75-1 cell lines served as negative controls. Each section was scanned at a high resolution (40X) using the Nanozoomer Digital Pathology equipment (Hamamatsu, Bridgewater, NJ).

The scoring system used for each antibody is listed in table (2.3) [364, 419, 420]. IHC staining for ER, PR, HER2 and Ki-67 were used as surrogate markers to classify breast cancer tumors into luminal A, luminal B, HER2-positive and triple-negative breast cancer. Luminal A was defined by being (ER positive, PR positive, HER2 negative and Ki-67 < 14%) and luminal B was defined by being (ER, PR and HER2-positive) or (ER positive, PR positive, HER2 negative and Ki-67 ≥ 14%). Triple-negative breast cancers were samples that lack expression of ER, PR and HER2. Samples with HER2-positive while ER and PR are negative were considered as HER2-positive subtype [421].

Immunofluorescence and Confocal Microscopy

Direct immunofluorescence staining was carried out using Ventana automated immunostainer. Dylights 550 and 650 fluorophores from Thermo Scientific Pierce protein biology products were used to label primary antibodies against p63 (p63 Ab-1; IgG clone 4A4, Thermo-scientific) and ANXA1 (ANXA1; IgG1 clone29/Annexin I, BD transduction laboratory). The slides were incubated with the labeled antibodies for six hours. The slides were next washed and DAPI was applied. Sections were cover slipped and kept at -20 °C.

Table 2.2: Antibodies used for immunohistochemistry

Antibody	Clone	Dilution	Provider	Retrieval method	Positive score
ER	SP1	RTU	Ventana	HIER pH 6	Allred score ≥ 3
PR	1E2	RTU	Ventana	HIER pH 6	Allred score ≥ 3
HER2	4B5	RTU	Ventana	HIER pH 6	3+
Ki-67	SP6	1/100	BioCare	HIER pH 6	>14%
ANXA1	Clone29	1/500	BD	HIER pH 9	0-4 low expression 6-12 high expression
FOXA1	Polyclonal	1/500	Abcam	HIER pH 6	0-4 low expression 6-12 high expression
VIM	3B4	1/100	Dako	HIER pH 6	0-4 low expression 5-12 high expression

ER, Estrogen receptor, PR, Progesterone receptor; HER2, Human epidermal growth factor receptor 2; ANXA1, Annexin A1, FOXA1, Forkhead box protein A1; VIM, Vimentin; RTU, ready-to-use; HIER, Heat-induced epitope retrieval.

Table 2.3: Immunohistochemical scoring system.

Marker	Subcellular localization	Criteria		Total Score	Results
		Cell proportion score (A)	Staining Intensity score (B)		
ESR1 and PR (Allred score)	Nuclei	0= no staining 1= <1% nuclear staining 2= 1-10% nuclear staining 3= 11-33% nuclear staining 4= 34-66% nuclear staining 5= 67-100% nuclear staining	0= No staining 1= weak staining 2= Moderate staining 3= Strong staining	A+B= 0-8	Negative= 0-2 Positive= 3-8
ANXA1, VIM, Ki-67, FOXA1	ANXA1→Nuclei, cytoplasm, membrane VIM→Cytoplasm Ki-67→ Nuclei FOXA1→ Nuclei	0= no staining 1= <1-10 % staining 2= >10-50% staining 3= >50- 70% staining 4= >70-100% staining	0= No staining 1= weak staining 2= Moderate staining 3= Strong staining	AxB=0-12	Low expression = 0-4 High expression = 6-12
HER2	Membrane	0= No staining was observed in invasive tumor cells 1+ = Weak incomplete membrane staining in any proportion of invasive tumor cells 2+ = Complete membrane staining that is either non uniform or weak in intensity but with obvious circumferential distribution in at least 10% of cells. 3+ = uniform intense membrane staining of > 30% of invasive tumor cells		0, 1+, 2+, 3+	0-1+ are negative 2+ is equivocal 3+ is positive

ESR1, Estrogen receptor, PR, Progesterone receptor; ANXA1, Annexin A1, FOXA1, Forkhead box protein A1; VIM, Vimentin; HER2, Human epidermal growth factor receptor 2

Fluorescent images were captured using Zeiss LSM 510 laser scanning confocal microscope equipped with 63x oil immersion objective lens. META detector was used for spectral separation of fluorochromes with overlapping spectra. Images were cropped, resized, and brightness and contrast over the whole image adjusted where necessary, using AxioVision and ImageJ softwares.

Statistical Analyses

All statistical analyses were carried out using different packages of language R. The distribution of ANXA1 in different histological grades and molecular subtypes is shown using box plots and bar charts. Non-parametric tests are used due to the nature of ordinal and categorical data. The overall relationship between ANXA1 scores and other characteristics (i.e. histological grades, and molecular subtypes) was evaluated by chi-square test. Correlation analysis for immunohistochemical expression levels was carried out using the Spearman's rho correlation coefficient. Statistical significance was considered, with a $p\text{-}value < 0.05$.

Results

ANXA1 is Highly Expressed in Basal-Like Breast Cancer *in silico* Analysis.

Comparison of mRNA levels within breast cancer molecular subtypes using the bc-GenExMiner database comprising over 3400 microarrays and the mRNA sequencing dataset from the Cancer Genome Atlas (TCGA) dataset that included 844 patients were carried out. In patient microarray data sets, 59% of the basal-like and 70% of normal-like breast cancer subtypes show high level of ANXA1 expression. In contrast, 21% of luminal A and only 7% of luminal B subtypes demonstrate high expression of ANXA1. For the HER2-positive subtype, the percentage of patients with high or low expression of ANXA1 is nearly identical (32% vs. 29%, Figure 2.1A). Similarly, in the patient sequenced transcriptomes, ANXA1 mRNA levels were found to be significantly overexpressed in the basal-like subtype when compared to both luminal A ($p\text{-}value < 0.001$) and B subtypes ($p\text{-}value < 0.001$) (Figure 2.1B).

ANXA1 is Overexpressed in Triple-Negative Breast Cancer Cell Lines.

To validate the results obtained from our *in silico* analysis, ANXA1 expression was studied in a panel of human breast cancer cell lines with varying degree of expression of ER, PR and HER2 (Table 2.4). ANXA1 expression was detected in all basal-like cell lines 184-B5, MCF10F, MDA-MB-468 and MDB-MB-231 [422]. Intense membrane staining was present in all triple-negative cell lines. However, cytoplasmic and nuclear staining intensity somewhat varied in these cell lines (Figure 2.2A-D). In sharp contrast, ANXA1 was absent from MCF7, ZR-75-1 and MDA-MB-361 cell lines (Figure 2.2A-F) which are all considered to belong to the luminal subtype categories [422, 423]. Taken together, our findings in breast cancer cell lines support our *in silico* analysis showing higher expression of ANXA1 in triple-negative breast cancer.

ANXA1 is Strictly Confined to Myoepithelial cells in Both Normal Breast Tissue and *In situ* Breast Cancers.

In normal breast tissue and *in situ* carcinoma, we found ANXA1 to be strongly expressed in myoepithelial cells and in a few adjacent stromal cells (Figure 2.3A&B). Myoepithelial cell identity was confirmed by their typical localization around the luminal cells, their characteristic flattened morphology and more directly by the co-expression of ANXA1 with the specific myoepithelial cell marker p63 using confocal microscopy. ANXA1 and p63 were co-localized in the very same myoepithelial cells either in normal mammary gland or in ductal carcinoma in situ (Figure 2.3C&D).

ANXA1 is Overexpressed in Subsets of Triple-Negative and HER2-Positive Breast Cancers

Variation in the levels of ANXA1 expression according to molecular subtypes of breast cancer was assessed. All samples collected from patients with luminal A and luminal B breast cancer lacked detectable expression of ANXA1 (Figure 2.4A&B). In

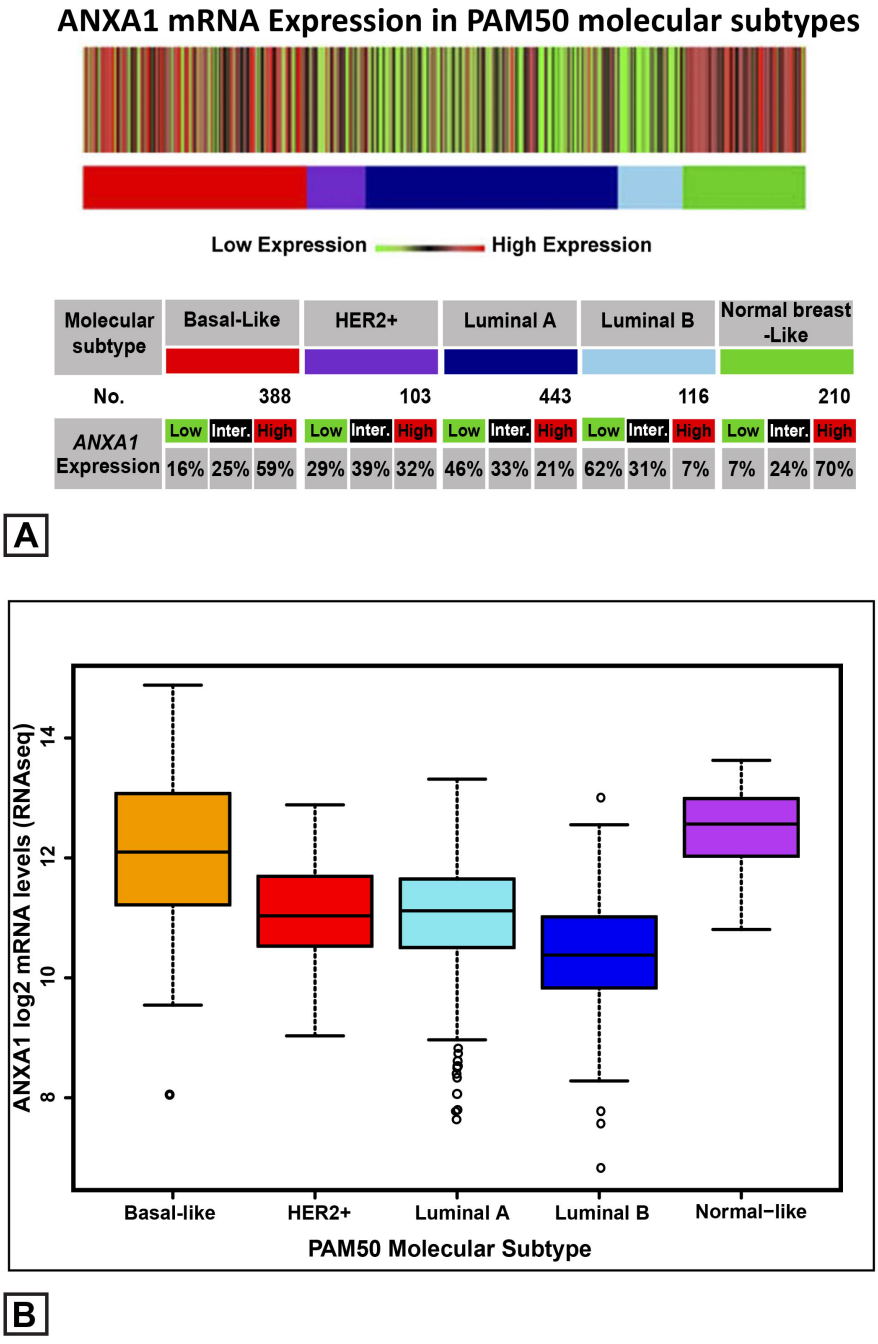


Figure 2.1: *In silico* analyses of ANXA1 expression in PAM50 molecular subtypes of breast cancer using A. The bc-GenExMiner database comprising over 3400 microarrays, B. TCGA mRNA sequencing dataset including 844 patients showing that ANXA1 is highly expressed in basal-like subtype. *** p-value < 0.001, Tukey's test.

Table 2.4: Expression of ER, PR, HER2 and ANXA1 in different types of breast cancer cell lines [422, 423]**.**

Cell Lines	ER	PR	HER2	Molecular Subtypes	ANXA1
184B5	-	-	-	Basal	+
MDA-MB-468	-	-	-	Basal	+
MCF10F	-	-	-	Basal	+
MDAMB231	-	-	-	Basal	+
ZR-75-1	+	+	-	Luminal	-
MDA-MB-361	+	-	+	Luminal	_
MCF7	+	+	-	Luminal	-

ER, Estrogen receptor, PR, progesterone receptor; HER2, Human epidermal growth factor receptor 2; ANXA1, Annexin A1.

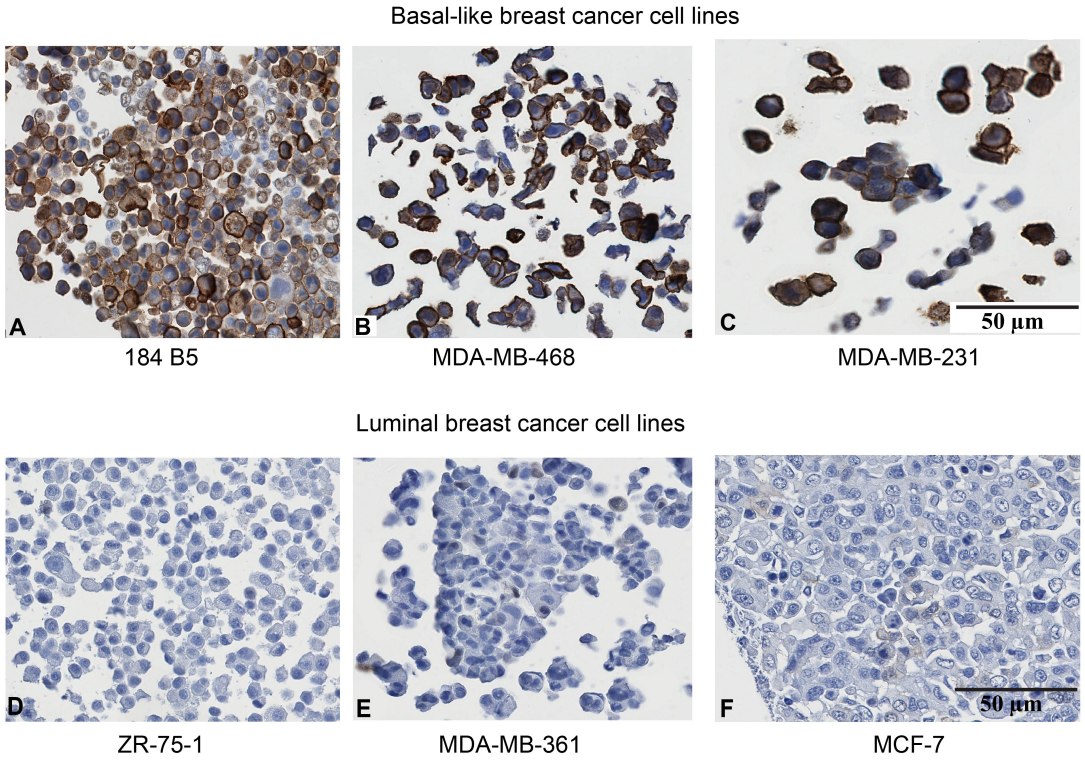


Figure 2.2. Expression of ANXA1 in different types of breast cancer cell lines. Strong ANXA1 expression in basal-like breast cancer cell lines (A-C). Lack of ANXA1 expression in luminal breast cancer cell lines (D-F). Magnification: 40X (A-F)

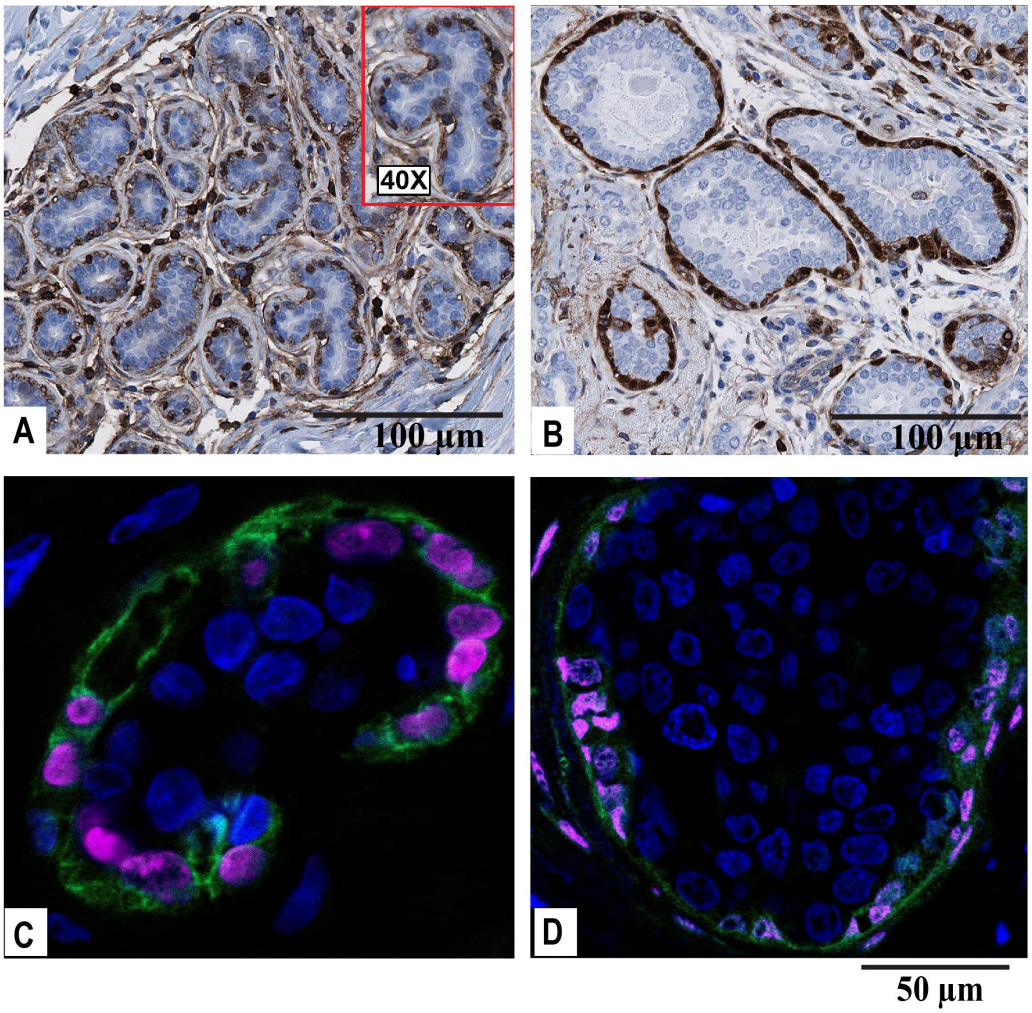


Figure 2.3. Expression of ANXA1 in normal breast tissue and *in situ* carcinoma using immunohistochemistry (A&B) and direct immunofluorescence (C&D). ANXA1 expression is restricted to myoepithelial cells and a few adjacent stromal cells. For figure 2.3 C&D, the green color mark ANXA1 and the purple mark the P63 Magnification: 20X (A & B), 40X inset in Figure A, 63X (C & D)

contrast, almost half of the samples collected from patients with triple-negative breast cancers had strikingly elevated levels of ANXA1 expression (Figure 2.4C&D). Moreover, the association of ANXA1 with basal-like breast cancer, which is considered to be a subset of triple-negative breast cancer, was further substantiated by a positive correlation

with basal cytokeratin CK5/6 ($r = 0.72$, *p-value* < 0.001) (data not included). Notably, most but not all HER2-positive samples (approx. 90%) were found to be negative for ANXA1 (Figure 2.4E,F&G).

Altered Pattern of ANXA1 Expression Correlates with a Subset of High-Grade Breast Cancers.

In Grade I, Grade II and 70.4 % of Grade III breast cancer, ANXA1 expression could not be detected in tumor cells (Figure 2.5A-C). However, 29.6% of Grade III breast cancer patients showed ectopic expression of ANXA1 in tumor cells (Figure 2.5D-G). Another key finding is the disordered cellular distribution of ANXA1 labeling in high-grade breast cancers. Whereas ANXA1 expression is normally restricted to the myoepithelial cells in a typically uniform fashion, the pattern observed in tumor cells markedly departed from that of normal cognate cells. Indeed, the staining pattern in tumors is at odds with that observed in normal myoepithelial cells. Some tumors exhibit ANXA1 staining mainly in the cytoplasm with only a few cells harboring nuclear staining. Other tumors demonstrate predominantly nuclear staining with only very faint cytoplasmic staining. On occasion, we have observed an intense nuclear and cytoplasmic staining. Lastly, strong membranous staining along with weak cytoplasmic staining was found in few high-grade breast cancers (Figure 2.5D-G). Taken together, our results indicate that not only ANXA1 is aberrantly accumulating in a significant minority of high grade tumor cells but also that there is an accompanying defect in its cellular translocation, distribution and compartmentalization

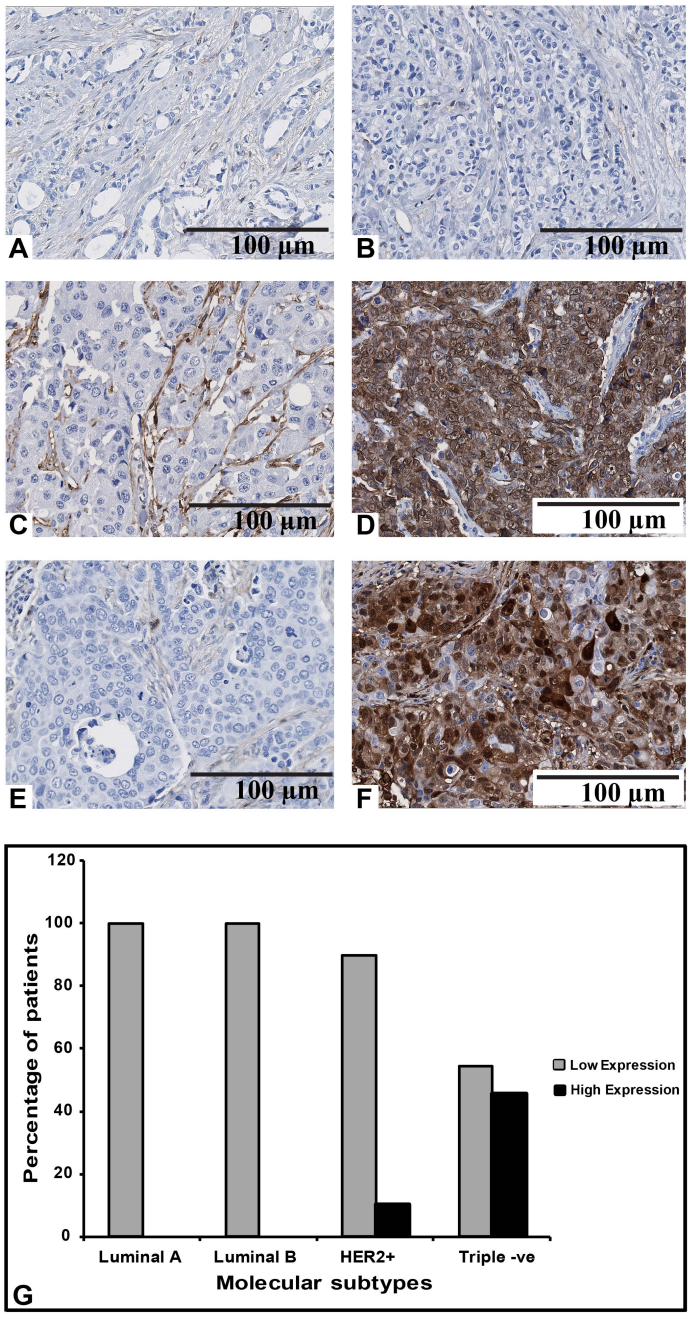


Figure 2.4. **Expression of ANXA1 in different molecular subtypes of breast cancer.** A. Luminal A, B. Luminal B. Luminal A and B tumor cells fail to express ANXA1. A subset of triple negative (C&D) and HER2-positive (E&F) tumors express high level of ANXA1 (D&F). G. Diagram showing the percentage of patients with high and low expression of ANXA1 in each molecular subtype of breast cancer Magnification: 20X (A-F)

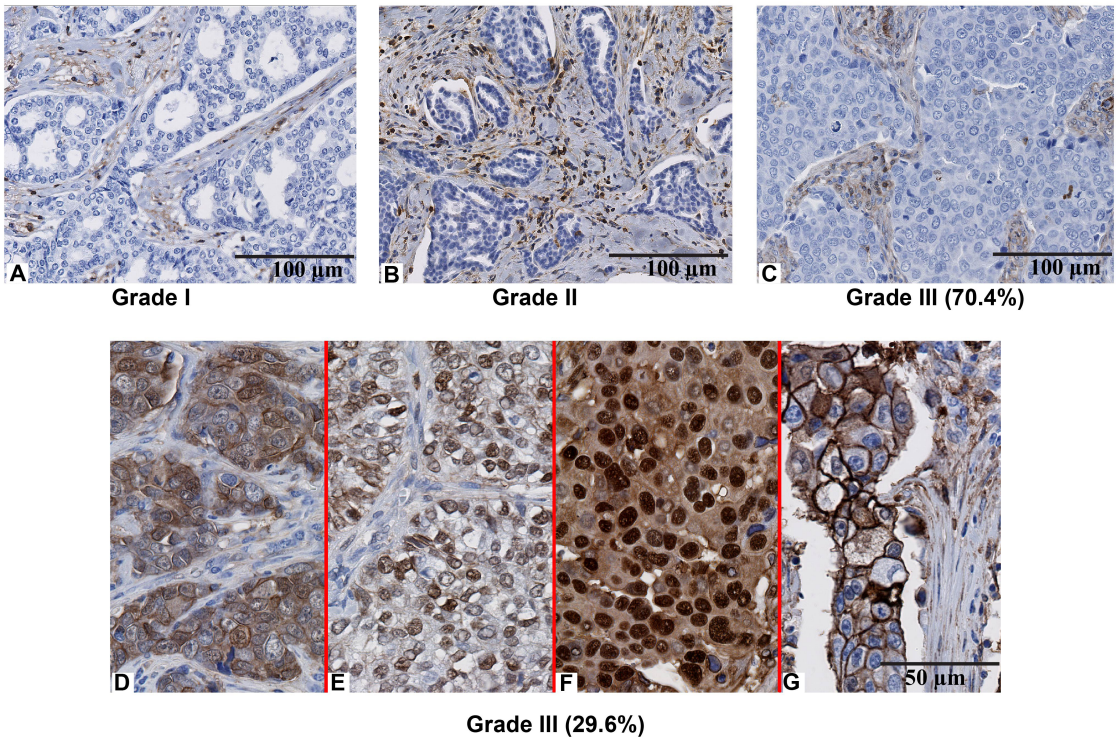


Figure 2.5. **ANXA1 expression in breast cancer of different histological grades.** 100% of Grade I & Grade II breast cancer and 70.4% of Grade III fail to express ANXA1 in the cancer cells while stromal cells show positive staining (A-C). However, 29.6% of Grade III breast cancers showed high ANXA1 expression along with varied subcellular localization (D-G). Magnification: 20X (A-C), 40X (D-G)

Positive Correlation Between ANXA1 and Vimentin and Negative Correlation Between ANXA1, ESR1 and FOXA1.

The RNA sequencing data derived from TCGA dataset [38, 416] was analysed using the MiSTIC visualization tool to identify genes that are positively or negatively correlated with ANXA1 in normal breast tissue (102 cases) and in breast cancer (756 patients). A positive correlation was observed between ANXA1 and VIM, one of the EMT markers in breast cancers (r =0.73). In the same group of tumors, ANXA1 is negatively correlated with several luminal markers such as FOXA1 (r =-0.66), ESR1 (r =-0.46), GATA3 (r =-0.47) and XBP1 (r = -0.47) (Figure 2.6). To substantiate these observations, we next assessed the mRNA levels of ANXA1, VIM, ESR1 and FOXA1 using publically available microarrays and mRNA sequencing breast cancer patients' datasets. The microarray included 51 breast cancer cell lines [416] and 247 patients [417] and mRNA sequencing included 844 patients. Similar correlations could be established using mRNA in sequenced transcriptome datasets strengthening the above mentioned association. Specifically, a positive correlation between ANXA1 and VIM was detected (r = 0.75 Figure 2.7A) while a negative correlation was established between ANXA1 and ESR1 (r = -0.47 Figure 2.7B) and between ANXA1 and FOXA1 (r = -0.46 Figure 2.7C).

In order to expand on the data obtained from the *in silico* analyses, we used additional sections of our human breast cancer TMAs to ascertain the expression of VIM, ESR1 and FOXA1 at the protein level. Again, a positive correlation between ANXA1 and VIM was confirmed (r = 0.48, *p-value* < 0.001). In triple-negative breast cancer, 62% of patients had a high expression of both proteins. In contrast, 100 % of luminal A patients and 95.6 % of luminal B had low expression of ANXA1 and VIM. Similarly, 77.3% of HER2-positive patients also had low expression of ANXA1 and VIM with only 4.6 % showing high expression of both proteins.

We were also able to confirm the negative correlation between ANXA1 and ESR1 (r =-0.45, *p-value* < 0.001) and between ANXA1 and FOXA1 (r =-0.51, *p-value* < 0.001) using TMAs. Specifically, in luminal A &B breast cancer, 100% of the patients had a high degree of expression of ESR1 and FOXA1 and no detectable levels of ANXA1 expression (Figure 2.8A). In triple-negative breast cancer where ESR1 and FOXA1 are negative, 45.5% of patients had high levels of ANXA1 (Figure 2.8B).

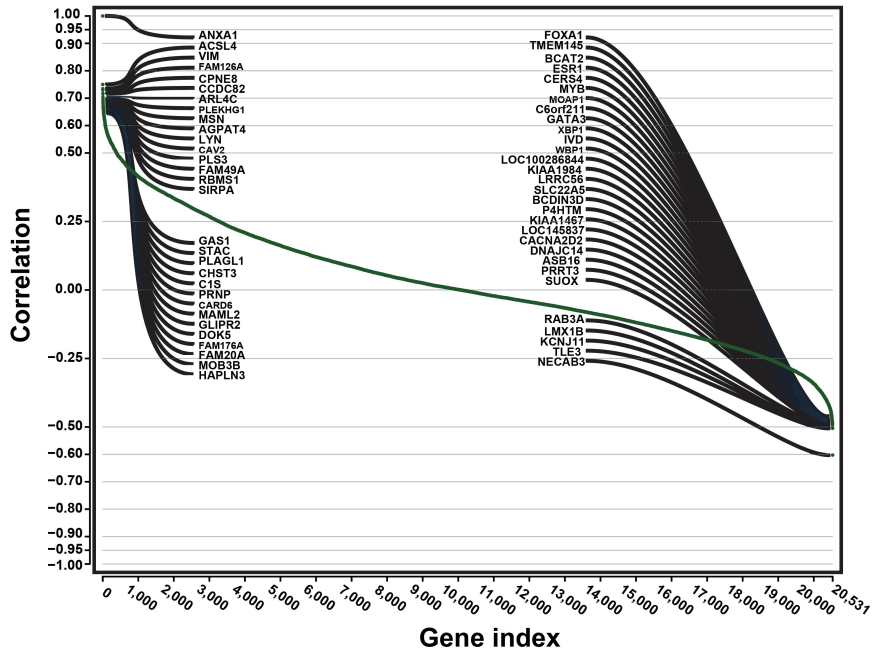


Figure 2.6. Genes that are positively (left side) or negatively correlated (right side) with ANXA1 in 756 breast cancer patients based on RNA-sequencing data derived from TCGA dataset.

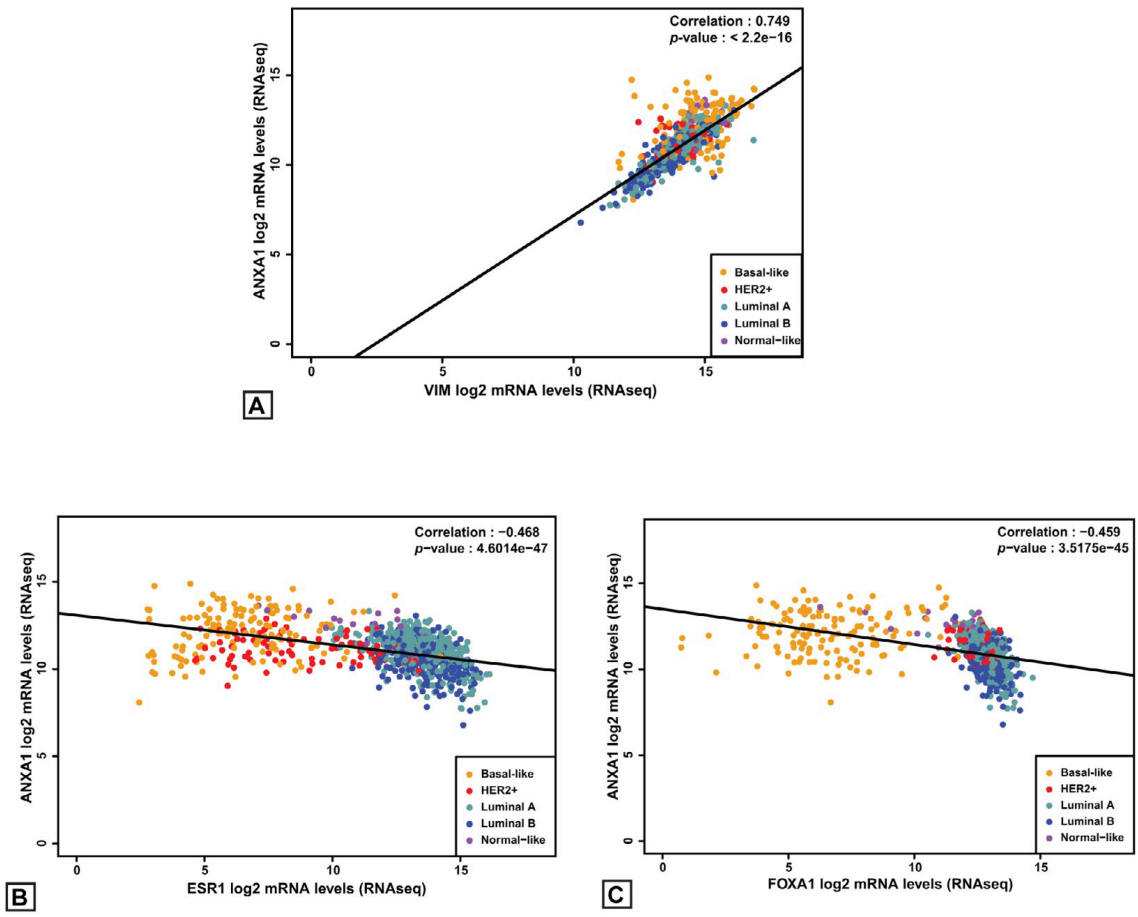


Figure 2.7. **Correlation between ANXA1 and VIM, ESR1 and FOXA1 using RNA sequencing data of 844 breast cancer patients.** There is positive correlation between ANXA1 and VIM. In contrast, ANXA1 is negatively correlated with both ESR1 and FOXA1.

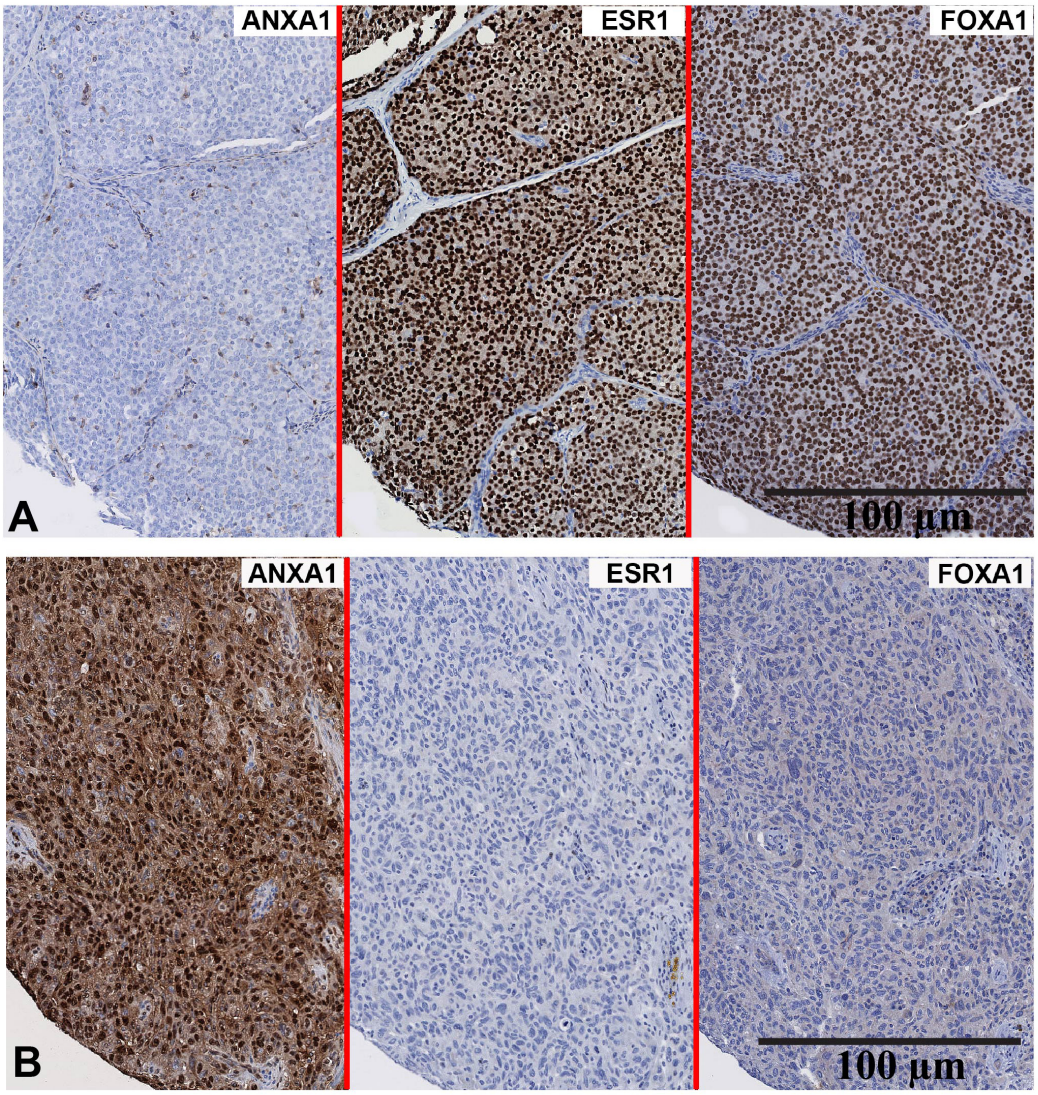


Figure 2.8. Correlation between ANXA1, ESR1 and FOXA1 in different molecular subtypes of breast cancer using immunohistochemistry and tissue microarray. A. Luminal breast cancers, B. Subset of triple negative breast cancer. Magnification: 10X

Discussion

ANXA1 belongs to a family of calcium and phospholipid binding proteins initially considered to have an anti-inflammatory function. However, many recent publications also stressed its role in key cellular processes such as proliferation, differentiation, cytoskeletal organization, cell migration and apoptosis. Although ANXA1 expression has been studied in different types of cancer, so far results have been rather conflicting, with reports in the literature describing variable levels of expression in normal and breast cancer tissue. This also holds true for breast cancer where the status of ANXA1 expression in normal mammary gland and breast cancers is currently unsettled. In the present work, we have investigated the expression of ANXA1 in human breast cancer tissues. We found that deregulated ANXA1 expression is observed only in high-grade breast cancers that comprise clinically aggressive tumors such as triple-negative and HER2-positive breast cancers supporting the view that ANXA1 is a valuable marker for a subset of aggressive breast tumors.

Using TMAs comprising normal breast tissue and breast carcinomas of various histological grades, we conclusively demonstrated that ANXA1 expression is confined to the myoepithelial cell layer and to a few adjacent stromal cells of normal breast tissue. Myoepithelial cell identity was confirmed by their typical localization around the luminal cells, their characteristic flattened morphology and more directly by the co-expression of ANXA1 with the specific myoepithelial cell marker p63 using confocal microscopy. This constitutive pattern of ANXA1 expression is also highly conserved in the myoepithelial cell layer surrounding the ducts and lobules of *in situ* carcinomas. Notably, there is a complete absence of ANXA1 in malignant epithelial cells where the normal regulatory pathways continue to operate. These results are consistent with those of Ang et al. who reported similar pattern of expression of ANXA1 in myoepithelial cells [359]. However, they are clearly at odds with the observations of Elshimali & Lui who found that ANXA1 is expressed in both epithelial and myoepithelial cells of normal breast tissue [424].

Since there is complete disappearance of the myoepithelial cells layer during stromal infiltration, the lack of ANXA1 positive myoepithelial cells was an expected

finding. This is consistent with the inability to detect ANXA1 in low-grade (Grade I) and intermediate-grade (Grade II) and 70.4 % of high-grade (Grade III) breast cancer. Strikingly however, 29.6% of Grade III breast cancer patients showed ectopic expression of ANXA1 in tumor cells. The morphology of tumor cells and the failure to express any of the other myoepithelial cell markers are inconsistent with the possibility that those cells represent a residual myoepithelial population. Although it is tempting to hypothesize that positive tumor cells are derived from a progenitor cell endowed with both luminal and myoepithelial features, this view nevertheless fails to explain the absence of ANXA1 positive tumor cells in both grade I and grade II tumors. In our opinion, ANXA1 expression in tumor cells more likely represents an abnormal, deregulated cellular process due to a failure of the checks and balances that operates under normal conditions in the breast tissue.

Our results on human breast cancer cell lines indicate that triple-negative cell lines expressed ANXA1 at significantly higher levels than those observed in both luminal A or B cell lines. This result is also consistent with the observations of Kang et al. that breast cancer cell lines with high expression of ANXA1 are invasive by nature, whereas lines with low or undetectable expression of ANXA1 had a poor capacity for migration [425]. Additionally, we found that enhanced expression of ANXA1 is a characteristic of a subset of triple-negative and, to a lower extent, HER2-positive breast cancers. It is also worth noting that we were able to relate ANXA1 expression to basal-like breast cancer as shown by a positive correlation with basal cytokeratin CK5/6. Again, this supports a similar finding by Yom et al. who demonstrated that ANXA1 expression was significantly correlated with unfavorable prognostic factors such as hormone receptor negativity, HER2-positive tumors and triple-negative breast cancer [379]. To our surprise though, Wang et al. found that low ANXA1 expression was significantly associated with advanced stage breast cancer and a worse overall survival when compared to patients with high ANXA1 expression [371]. The fact that an enhanced expression of ANXA1 was found in the so-called normal-like breast cancer subtype supports the view that this molecular category may simply reflect the high degree of contaminating normal breast tissue and hence residual myoepithelial cells in the tumor samples [426]. Future

experiments with preclinical models will help solving this issue and/or this apparent contradiction

Concerning the putative role of ANXA1 in the EMT, in this study we found a positive correlation between ANXA1 and VIM, one of EMT markers. Supporting data came from both the mRNA level using *in silico* analysis and from ANXA1 cell labeling on the TMAs. This may suggest that ANXA1 exerts a positive effect in EMT, consistent with reports indicating that ANXA1 promotes metastasis formation in basal-like breast cancer cells by enhancing TGFβ/smad signaling and actin reorganization [375]. Likewise, ANXA1-null mice were found to develop fewer metastases than wild type littermates [414]. However tantalizing this hypothesis might be, it is severely challenged in view of the findings by Maschler et al. who demonstrated that forced ANXA1 expression in metastatic mouse and human mammary carcinoma cells reversed EMT and abolished metastasis [376]. Clearly, more work is needed to sort out the exact role of ANXA1 in tumor progression and metastasis. The use of a comprehensive panel of EMT markers on human breast tumors expressing ANXA1 could help clarifying the issue. Of note, MSN and LYN that are known to play roles in EMT [427, 428] have also positive correlation with ANXA1 (For MSN r =0.69, for LYN r =0.54). (Figure 2.6)

A negative correlation between ANXA1 and both ESR1 and FOXA1 has been uncovered using *in silico* study. This was further substantiated using a large cohort of breast cancer patients tissues and cell lines. 100% of patients with luminal breast cancer subtypes had positive expression of ESR1 and FOXA1 while they completely lack ANXA1 expression. In contrast, 43% of triple negative breast cancers that are ESR1 and FOXA1 negative, display high levels of ANXA1. Recently, FOXA1 was reported to be necessary not only for maintaining luminal-specific gene expression, but also for repressing numerous genes specific to basal breast cancer cells including ANXA1 [429]. Interestingly, our *in silico* study confirmed that ANXA1 promoter contains FOXA1 responsive element (two regions within 5kb of ANXA1 transcription start site) but none for ESR1. Given the suppressive roles of FOXA1 [430-432], it would be of great interest to determine how FOXA1 suppresses ANXA1 expression.

Lastly, one has to remind that ANXA1 is only one member of a broad family of annexins many of which have been reported to be directly or indirectly involved in breast

cancer tumorigenesis such as Annexin A2, A4, A5, A6, A7, A8 [38, 416, 433, 434]. At present, the degree of functional overlap and crosstalk between each annexin variant has not been addressed.

To conclude, ANXA1 expression is restricted to high-grade breast cancers, mostly triple-negative breast tumors that are notoriously known for their aggressive clinical behavior. We propose that ANXA1 is deregulated in a significant proportion of high histological grade breast cancers, thereby underlining the complexity and heterogeneity of breast cancers especially those that belong to the triple-negative category. Developing new biomarkers will help dissect out subsets of triple-negative tumors and facilitate the identification and further characterization of unique and specific regulatory pathways paving the way to targeted therapy. A better understanding of the complex cross-regulatory networks between ANXA1, Vimentin, ESR1 and FOXA1 signaling pathways is clearly needed to pinpoint and identify subsets of tumors with non-overlapping and unique mechanisms leading to tumor progression and metastasis.

CHAPTER III

MMP-9 expression varies according to molecular subtypes of breast cancer

Einas M Yousef[1], Muhammad R Tahir[2], Yves St-Pierre[3], Louis A Gaboury[1&4§]

[1]Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, Canada, [2]Le centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Canada, [3]INRS-Armand-Frappier, Université du Québec, Laval, Montréal, Canada, [4]Department of Pathology and Cell Biology, Faculty of Medicine, Université de Montréal, Montréal, Canada
[§]Corresponding author

Published in BMC Cancer, August 2014

Authors' contributions

EY, YSP and LG are responsible for the study design.

EY performed the experiments and collected the data.

EY performed the *in silico* analysis.

EY and LG scored the IHC reactions.

EY prepared all the diagrams and pictures.

MT carried out the statistical analysis.

EY, YSP and LG participated in the data analysis and interpretation.

All authors drafted, read and approved the final manuscript.

Abstract

Background

In 2014, breast cancer remains a major cause of mortality worldwide mostly due to tumor relapse and metastasis. There is currently a great interest in identifying cancer biomarkers and signalling pathways mechanistically related to breast cancer progression. Matrix metalloproteinase-9 (MMP-9) is a member of matrix degrading enzymes involved in cancer development, invasion and metastasis. Our objective was to investigate MMP-9 expression in normal human breast tissue and to compare it to that of breast cancer of various histological grades and molecular subtypes. We also sought to correlate MMP-9 expression with the incidence of metastasis, survival rates and relapse in breast cancer patients.

Methods

MMP-9 was first studied using *in silico* analysis on available DNA microarray and RNA sequencing data of human breast cancer tissues and human breast cancer cell lines. We next ascertained MMP-9 expression in both normal breast tissue and in human breast carcinoma tissue microarrays.

Results

Significant increase in MMP-9 expression was found in breast cancer cells when compared to normal breast tissue. A positive correlation could also be established between elevated levels of MMP-9 and breast cancer of high histological grade. Furthermore, our results indicate that not only MMP-9 is differentially expressed between each molecular subset but also; more importantly MMP-9 overexpression revealed itself as a startling feature of triple-negative and HER2-positive breast cancers. Lastly, the clinical relevance of MMP-9 overexpression is strongly supported by its significant association with a higher incidence of metastasis and relapse.

Conclusions

Differential expression of MMP-9 reflects the extent of cellular differentiation in breast cancer cells and is closely related to the most aggressive subtypes of breast cancer. Hence, MMP-9 is a promising prognostic biomarker of high-grade breast cancer. In our opinion, MMP-9 expression could help segregate subsets of aggressive breast cancer into clinically meaningful subtypes

Keywords: MMP-9, human breast cancers, metastasis, *in silico* analysis, tissue microarrays.

Competing interests

The authors declare that they have no competing interests.

Introduction

Breast cancer is the most common malignancy and the second leading cause of cancer-related death after lung cancer among women in the United States and Europe [435]. Due to major advances in screening and early diagnostic procedures, most breast cancer patients are diagnosed at an early stage. However, 6% to 10% of patients still present with metastatic breast cancer at the time of diagnosis; for those patients, relapses tend to occur earlier and survival rates are shortened [436]. Cancer metastasis is considered to develop in a step-wise fashion leading to the acquisition of new capabilities by tumor cells helping them to thrive and evade natural barriers [437]. Cancer cells detach themselves from the primary tumor, migrate and invade surrounding tissues, enter the vasculature, circulate throughout the body and eventually reach secondary sites where they extravasate, and populate distant organs [438].

Degradation of the extracellular matrix (ECM) is thought to be a crucial step in the formation of tumor metastasis. Multiple proteolytic enzymes such as plasmin, cathepsins, and matrix metalloproteinases (MMPs) are known to degrade ECM [439]. Matrix metalloproteinase-9 (MMP-9) is a zinc-dependent peptidase that belongs to the gelatinase subfamily of MMPs. It is excreted as an inactive pro-enzyme that undergoes activation upon cleavage by different types of extracellular proteases [381]. MMP-9 activity is thought to be regulated by different biochemical stimulators such as growth factors and cytokines whose expression appear to modulate intracellular signaling pathways [382]. MMP-9 has the ability to degrade denaturated collagens which have been first cleaved by various collagenases such as MMP-1, MMP-8 and MMP-13 [440, 441]. In addition, MMP-9 degrades type IV collagen which is the main component of the basement membrane [442]. It exerts different roles in the dissemination process such as tumor invasion, tumor-induced angiogenesis, and immunomodulation of the tumor microenvironment. In addition, MMP-9 is instrumental in creating so-called premetastatic niches that foster colonization of distant organs [443]. Elevated tissue levels of MMP-9 are also associated with invasion, metastasis and poor prognosis in different types of cancer including cervical [395], colorectal [396], ovarian [397] and breast cancer [398]. Furthermore elevated levels of MMP-9 in the serum and urine have

also been found to be associated with metastasis and poor prognosis in a diversity of cancers [399].

Our goal was to assess the potential clinical usefulness of MMP-9 as a prognostic biomarker of breast cancer. To achieve that aim, we first studied *MMP-9* mRNA expression using *in silico* analysis on available DNA microarray and RNA sequencing data of human breast cancer tissues and breast cancer cell lines. We next evaluated MMP-9 expression at the protein level using immunohistochemical analyses on tissue microarrays containing both normal and neoplastic breast tissues. Our data were next correlated with patients' outcome specifically looking at the incidence of metastases, relapse and overall survival. Our results indicate that MMP-9 is not only differentially expressed in different molecular breast cancer subtypes but also overexpressed in triple-negative and HER2-positive breast cancers. Overexpression of MMP-9 tightly correlates with a higher incidence of metastasis and relapse. Taken together, our data indicate that differential expression of MMP-9 reflects the degree of differentiation of breast cancer cells and that its overexpression tightly correlates with the most aggressive subtypes of breast cancers. Hence, MMP-9 is a potentially useful biomarker of aggressive and metastatic subtypes of breast cancer.

Material and Methods

In silico Analysis

The web application bc-GenExMiner [415] was used for correlation analysis of MMP-9 gene expression on a dataset comprising over 3,063 microarrays. However, only 1210 patients could be correctly assigned precisely to each molecular subtype. The "aov" and "TukeyHSD" functions were carried out to compare the mRNA levels within each breast cancer molecular subtypes. The ANOVA was applied to check for an overall difference of expression levels between each molecular subtypes. The Tukey multiple comparisons of means were used to test for a significant difference between two subtypes (e.g. luminal A vs. basal). For both tests, a p-value < 0.05 was considered significant. The

mRNA level of MMP-9 in 51 breast cancer cell lines were also studied using publically available microarrays and mRNA sequencing breast cancer cell line datasets [416].

Patients and Tissue Samples

A retrospective study was carried out using a cohort of 300 female breast cancer patients comprising tumors of different histological grades. Archived formalin-fixed, paraffin-embedded (FFPE) samples containing tumor tissues were collected for the study. Tumor grades were confirmed using the Modified Scarff-Bloom-Richardson-Elston-Ellis grading system (SBR-EE) [444]. A complete set of follow-up data including the onset of metastasis and relapse were acquired. We also obtained 19 normal breast tissues from healthy women undergoing plastic surgery to serve as internal controls. Benign breast conditions such as mammary fibroadenoma and myofibroblastoma were included as negative controls [403]. In addition, a number of extraneous tissues such as colon, thyroid and placenta were included in each TMA. All samples were obtained from *Centre Hospitalier de l'Université de Montréal* (CHUM) after granting the approval of the research ethical committee (Comité d'éthique de la recherche du CHUM CENTRE DE RECHERCHE, Approval No. SL 05.019).

Tissue Microarray (TMA)

Sections (4 μm) from each paraffin block were stained with hematoxylin and eosin (H&E) and examined by two independent pathologists. Core punches, 1mm in diameter, were drilled from representative areas contained within each FFPE tumor blocks. Each core was realigned in duplicate or triplicate into recipient blocks according to the intended design of the map using a Manual Tissue Arrayer I (Beecher Instruments). Blocks were next inverted and incubated overnight in the oven over a glass slide. TMA blocks were allowed to cool until they could easily detach from the glass slide. Tissue sections from each TMA were prepared and one slide from each block was stained with H&E to review the diagnoses and histological grades on all tissue samples. Additional representative sections from each block were submitted to automated immunohistochemical (IHC) staining.

Immunohistochemistry

Immunohistochemical assays were performed on FFPE tissues obtained from each TMAs. These assays were carried out according to manufacturer recommendations on an automated immunostainer (Discovery XT system, Ventana Medical Systems, Tucson, AZ). Immunohistochemical analysis of MMP-9 (polyclonal; ab38898, dilution 1/100, no pretreatment, Abcam, Canada) was carried out to detect both the pro- and the active form of MMP-9 [445]. In addition, immunohistochemical analysis of estrogen receptor (ER; monoclonal, clone SP1, RTU, sCC1, Ventana Medical Systems), progestrone receptor (PR; monoclonal, clone 1E2, RTU, sCC1, Ventana Medical Systems), HER2 (monoclonal, clone 4B5, RTU, sCC1, Ventana Medical Systems), Ki-67 (monoclonal, clone SP6, dilution 1/100, pretreated sCC1, BioCare medical) were used as surrogate markers of breast cancer molecular subtypes [421]. Antigen retrieval was performed with proprietary reagents followed by incubation with the primary antibody. Sections were then incubated with a specific secondary biotinylated antibody for 32 minutes. Streptavidin horseradish peroxidase, and 3,3-diaminobenzidine were used according to the manufacturer's instructions (DABmap detection kit, Ventana Medical Systems). Sections were next counterstained with Gill's hematoxylin and sodium bicarbonate. Finally, each slide was scanned at high resolution (40X) using the Nanozoomer Digital Pathology equipment (Hamamatsu, Bridgewater, NJ). Two independent pathologists reviewed all stained sections on two separate occasions.

Estrogen receptor (ER) and progesterone receptor (PR) status were scored using Allred's method. In brief, the sum of the proportion and average intensity scores of positive tumor cells were calculated and results displayed on a scale ranging from 0 to 8. The cutoff point used to differentiate between positive and negative samples were as follows: tumors with Allred scores ≥ 3 (corresponding to as few as 1% to 10% weakly positive cells) were considered to be positive. Those tumors that had Allred score of less than 3 were considered to be negative. HER2 overexpression was carried out according to the College of American Pathologists (CAP)-approved scoring system as follows: no immunostaining or membrane staining which is incomplete or barely perceptible within $\leq$ 10% of the invasive tumor cells $\rightarrow$ 0; incomplete membrane or barely perceptible staining within >10% of invasive tumor cells $\rightarrow$ 1+; circumferential membrane staining

that is incomplete and/or weak/moderate within >10% of the invasive tumor cells or complete membranous staining that is intense within ≤ 10% of the invasive tumor cells → 2+ and circumferential membranous staining that is complete and intense → 3+ [420]. Scoring of MMP-9 and Ki-67 expression on each core was carried out using a two tier scoring system. The first parameter corresponds to the percentage of immunoreactive cells also known as the quantity score (QS). QS was estimated as follows (no staining was scored as 0, 1-10% of cells with positive staining were scored as 1, >10- 50% as 2, >50-70% as 3, and >70-100% as 4). We next assessed the second parameter (staining intensity score), which was rated as follows: No staining → 0, weak staining →1, moderate staining → 2, and strong staining → 3. The product of multiplying the quantity and the staining intensity scores represents the total IHC score that ranges from 0 to 12 [364, 446]. IHC scores of 0 to 4 were considered to represent low levels of expression while score from >4 to 12 were considered as high levels of expression.

IHC staining for ER, PR, HER2 and Ki-67 were used as surrogate markers to classify breast cancer tumors into luminal A, luminal B, HER-2 positive and triple-negative breast cancer. Luminal A was defined as being (ER positive, PR positive, HER2-negative and Ki-67 < 14%), luminal B was defined as being either (ER, PR, HER-2 positive) or (ER positive, PR positive, HER2-negative and Ki-67 ≥ 14%). Triple negative breast cancers consisted of tumors that lack expression of ER, PR and HER2. HER2- positive tumors that failed to express either ER or PR were considered to belong to the HER2-positive subtype [421].

Statistical Analyses

All statistical analyses were carried out using different packages of the R language (http://www.R-project.org/). The distribution of MMP-9 among different molecular subtypes is depicted using bar charts. Non-parametric tests were used due to the nature of ordinal and categorical data. The overall relationship between MMP-9 scores and molecular subtypes was evaluated using the chi-square test. Correlation analysis for immunohistochemical expression levels was carried out using the Spearman's rho correlation coefficient. Chi-square test was realized with Yates' continuity correction and a two-sided Fisher exact test was performed to analyze

metastasis. Kaplan-Meier plot was drawn to show the overall survival for low-level and high-level expression of MMP-9. Statistical significance was considered, with a *p-value* less than 0.05. Univariate and multivariate logistic regression were used to identify the significant factors among histological grades, histological subtypes, molecular subtypes, metastasis and age that affect the level of MMP-9 expression. The results were interpreted in terms of odds ratio (OR). Univariate and multivariate Cox models were used in survival analysis and the results were interpreted in terms of relative risk (RR). Statistical significance was determined by the confidence interval (CI). Only CI that does not include 1 are considered significant.

Results

In silico analysis: MMP-9 is overexpressed in basal-like and HER2-positive breast cancers.

The web application bc-GenExMiner [415] was used to compare the mRNA levels within each breast cancer molecular subtype on a dataset comprising 1210 microarrays. In brief, the gene expression data is given for those patients that could be assigned to a certain molecular subtype (robust classifications for 1210 patients). In figure 3.1, the table indicates for each subtype the proportion of patients with low, intermediate, and high gene expression. Gene expression values were being beforehand split in order to form three equal groups. This means that "high expression" is the 1/3 of the patients with highest expression of *MMP-9* and "low expression" is the lower 1/3 of the patients. As depicted in Figure 3.1, 57% of basal-like and 50% of HER2-positive breast cancer patients expressed high levels of *MMP-9*. In comparison, only 12% of those subtypes had a reduced expression of *MMP-9.* In sharp contrast, only 16% of the luminal A breast cancer subtype demonstrate increased expression of *MMP-9*. Data from the luminal B subtype indicate that 36% of patients have high levels of *MMP-9* expression while approximately 30% maintained low levels of *MMP-9*.

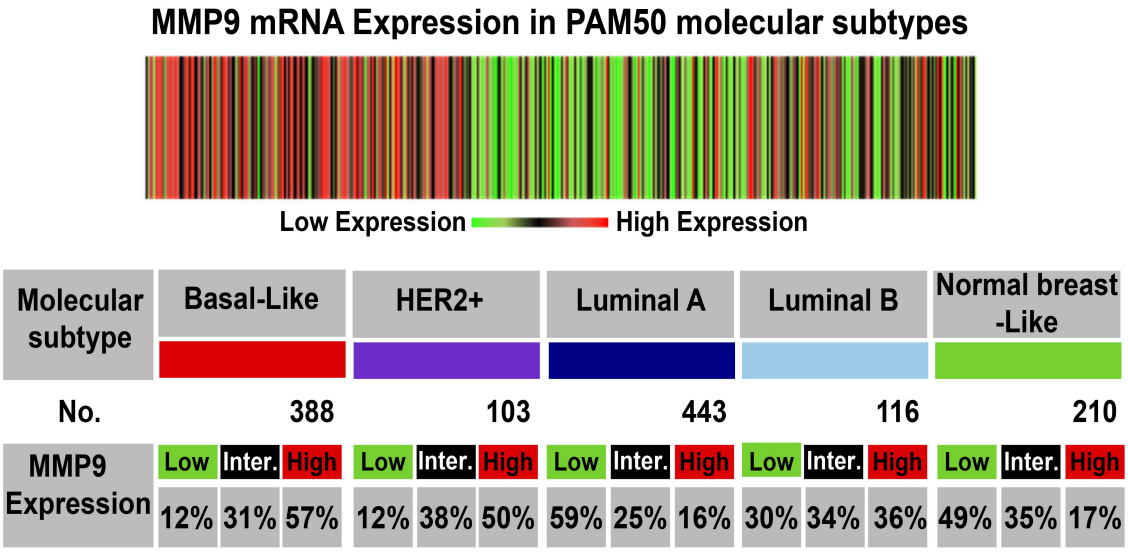


Figure 3.1: *In silico* analysis of *MMP-9* mRNA expression in breast cancer subtypes. The heat map and table are produced from the bc-GenExMiner database v3.0 showing the expression of *MMP-9* at mRNA level in different molecular subtypes of breast cancer as determined by PAM50. Overexpression of *MMP-9* is associated with basal-like and HER2-positive breast cancers. The "aov" and "TukeyHSD" functions were carried out to compare the mRNA levels within each breast cancer molecular subtypes.

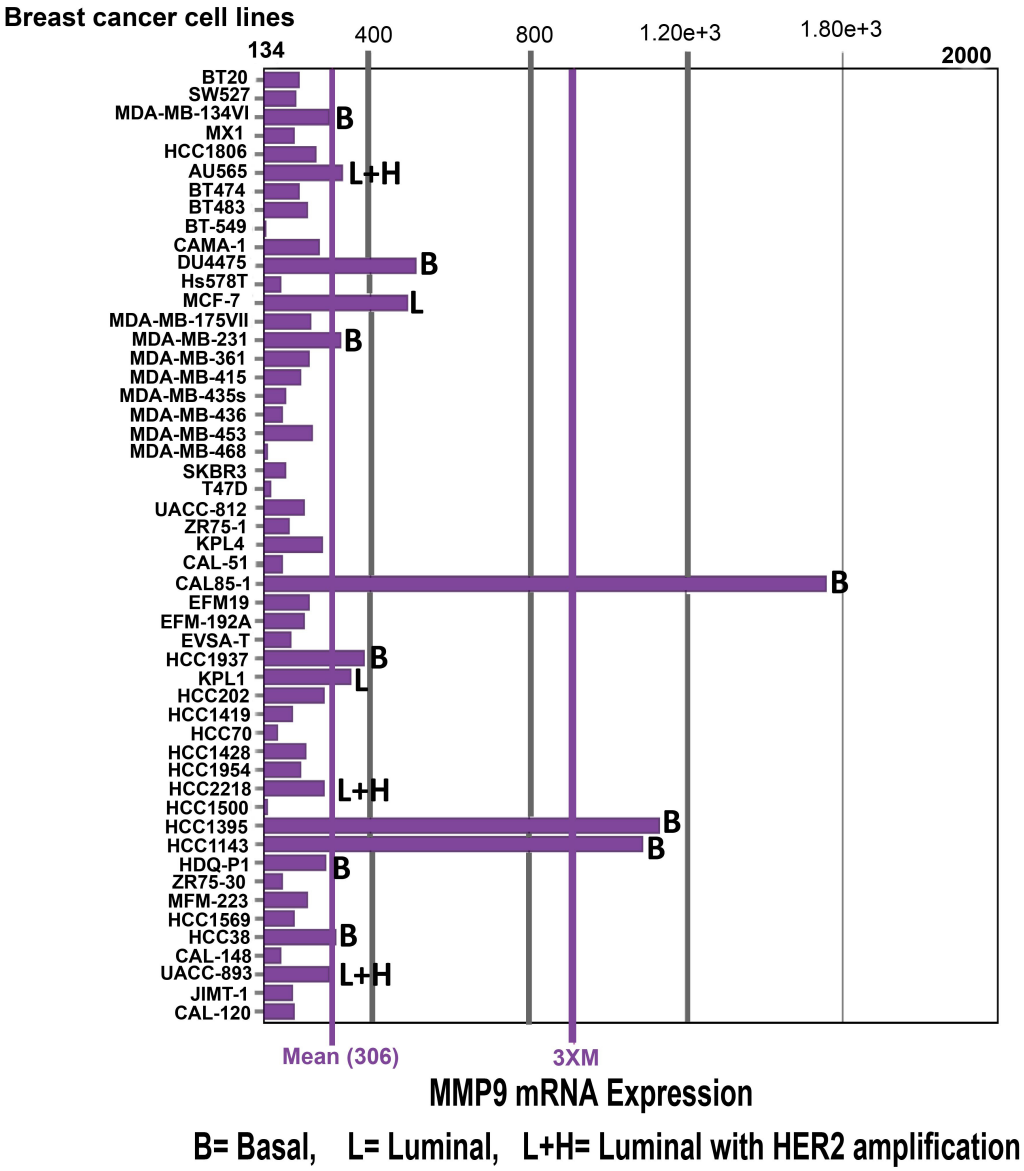


Figure 3.2: Expression of *MMP-9* mRNA in human breast cancer cell lines.

In silico analysis showing elevated *MMP-9* mRNA expression levels in basal-like breast cancer cell lines (e.g. CAL85-1, HCC1395, HCC1143, DU4475, HCC1937, MDA-MB-231 and HCC38). Luminal breast cancer cell lines with HER2 amplification also display stronger *MMP-9* mRNA expression (AU565, UAA-893 and HCC2218). MCF7 and KPL1 cell lines are the only luminal cell lines with mildly elevated *MMP-9* mRNA expression. (B= basal, L= luminal, L+H= Luminal with HER2 amplification)

To expand on the results obtained from the microarray datasets, we investigated mRNA expression of *MMP-9* in 51 breast cancer cell lines of different molecular subtypes [447-449] using publically available microarrays and mRNA sequencing breast cancer cell line datasets [416]. As shown in Figure 3.2, overexpression of *MMP-9* was present in basal-like breast cancer cell lines CAL85-1, HCC1395, HCC1143, DU4475, HCC1937, MDA-MB-231 [450] and HCC38. Interestingly, many luminal breast cancer cell lines known to have HER2 gene amplification (AU565, UAA-893 and HCC2218) also exhibited high levels of *MMP-9* expression. Notably, MCF7 and KPL1 cell lines were the only luminal cell lines that revealed a modest increase in *MMP-9* expression above baseline levels [451].

MMP-9 expression is markedly decreased or absent in normal human breast tissue.

Optimization of MMP-9 immune reactivity was a prerequisite to validating the specificity of the IHC reaction. In accordance with the Human Protein Atlas [95] and a review of the literature, human colorectal carcinoma was used as a positive control to assess the levels of MMP-9 expression in human cancers [452]. Our results are in complete agreement with this prediction as shown by the strong cytoplasmic labeling observed in colorectal carcinoma cells (Figure 3.3A). Additional adjacent sections from the same colonic tumor incubated with a non-immune serum containing IgG (same isotype/ same species) remained entirely negative. Of note, all subsequent steps of the immunostaining reaction such as addition of the secondary antibody and the revealing reaction were carried out in a strictly identical fashion (Figure 3.3B). We also thought fit to include benign breast lesions such as myofibroblastoma (Figure 3.3C) and fibroadenoma (Figure 3.3D) as negative controls [403]. Again, no immune reactivity could be detected after the successive addition of MMP-9 primary antibody, secondary antibody and chromogen.

Once all immunostaining conditions were satisfactorily established, we carried out IHC reactions on TMAs comprising both normal and neoplastic breast tissues. Our results indicate that 74% of normal breast tissues fail to express any MMP-9 reactivity in the luminal, myoepithelial cells and stromal cells surrounding normal breast ducts (Figure 3.4A). However, in a minority of normal breast tissues (26%) MMP-9 was faintly

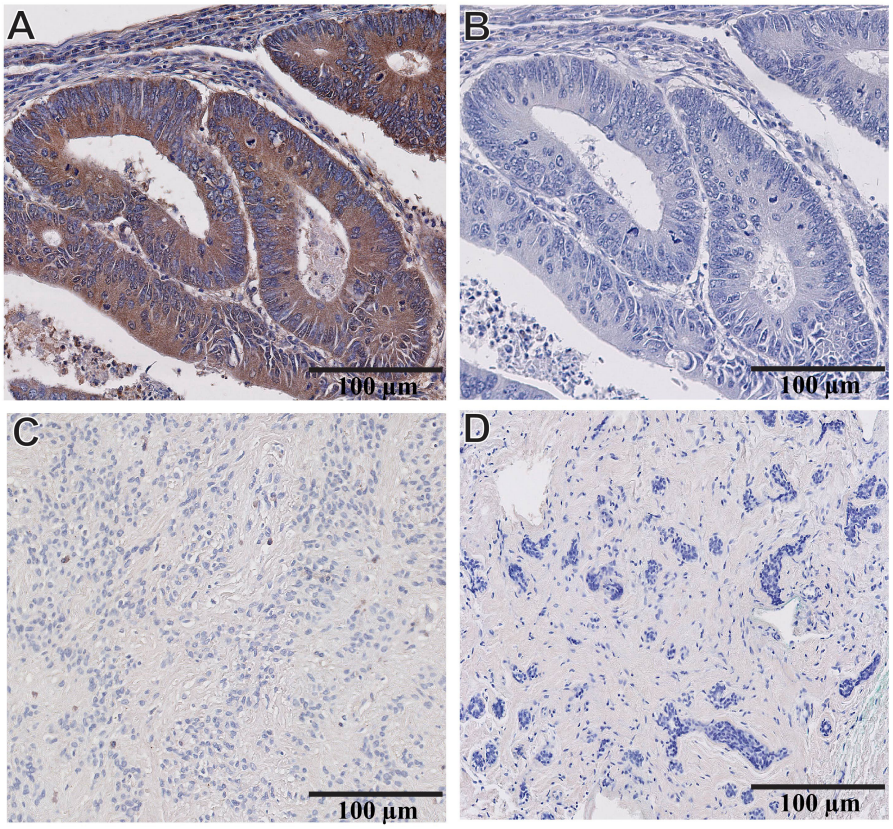


Figure 3.3: Validation of MMP-9 antibody specificity for IHC studies.

(**A)** Human colorectal carcinoma with intense cytoplasmic labeling of the cancer cells after incubating the section with MMP-9 primary antibody. (**B)** Adjacent section from the same colorectal tumor incubated with a non-immune serum that contains IgG (same isotype/ same species) showing complete lack of expression of MMP-9. (**C)** Benign myofibroblastoma of breast tissue and (**D)** Benign breast fibroadenoma do not express MMP-9. Magnification 20X (A-D)

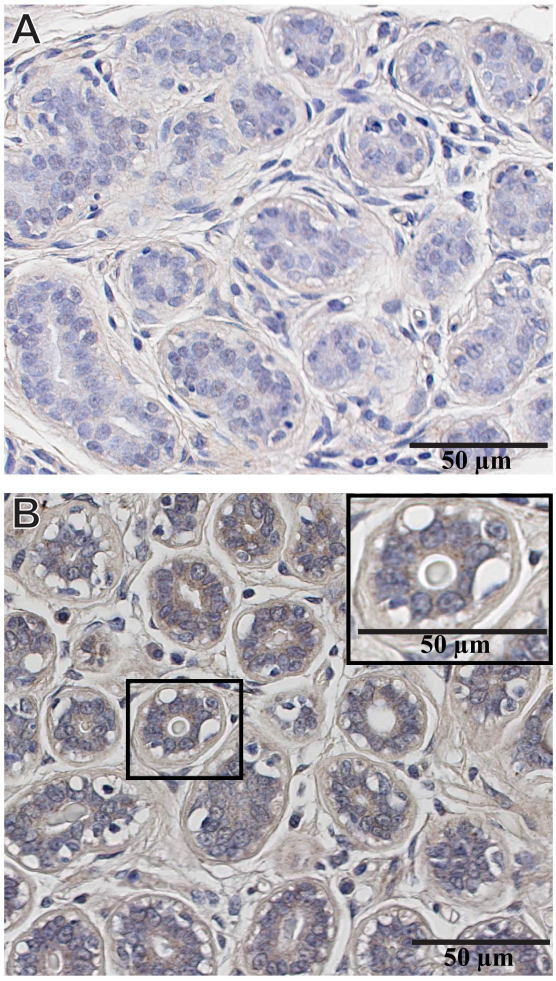


Figure 3.4: Expression of MMP-9 in normal breast tissue.

(**A)** Normal breast lobule lacking MMP-9 expression in both luminal and myoepithelial cells. Adjacent stromal cells also fail to express MMP-9 (74% of the patients). (**B)** Normal breast tissue exhibiting faint expression of MMP-9 in the cytoplasm of luminal cells, myoepithelial cells and in a few stromal cells surrounding normal breast acini. **A & B** are two distinct normal breast tissue from the same TMA incubated with anti-MMP-9 antibody. Magnification 40X (A&B), 63X inset in Figure 3.4B.

expressed and restricted to the cytoplasm of luminal, myoepithelial and a few adjacent stromal cells (Figure 3.4B). MMP-9 did not label either the nucleus or the cell membrane of any of these cells. Notably, the level of MMP-9 expression in the luminal cells consistently exceeded that present in the adjacent stromal cells.

Elevated levels of MMP-9 are present in carcinoma cells of triple-negative, HER2-positive tumors and nodal metastases.

Next we aimed to validate the results obtained from the *in silico* analysis on human breast tissue. We studied the expression of MMP-9 at the protein level and assessed the cellular and subcellular localization of MMP-9. MMP-9 expression was evaluated in 300 human tumor tissues representative of each molecular subtypes of breast cancer whose definition was based on the use of the following surrogate markers: ER, PR, HER2 and Ki-67 [421]. As shown in figure 3.5A, only 33.3% of luminal A ($p = 0.05$) and 43.3% of luminal B ($p < 0.01$) expressed elevated levels of MMP-9. In contrast, high levels of MMP-9 expression were found in 87.9% of HER2-positive and 79.4% of triple-negative breast cancer when compared to normal ($p < 0.001$). Low levels of MMP-9 expression were detected in the cytoplasm of cancer cells in both luminal A and B breast tumors. Indigenous stromal cells surrounding cancer cells in luminal A and B revealed only faint levels of MMP-9 expression (Figure 3.5B and C). On the other hand, elevated levels of MMP-9 expression were detected in the stroma surrounding cancer cells in both triple-negative and HER2-positive breast cancer. Nevertheless, the level of MMP-9 in the cytoplasm of cancer cells always exceeded that found in adjacent stromal cells (Figure 3.5D and E). Furthermore, when MMP-9 levels were evaluated in the cytoplasm of carcinoma cells present in 13 metastatic lymph nodes, it was found that tumor cells in 100% of our samples displayed elevated levels of MMP-9 whereas the surrounding lymphocytic and stromal cells failed to express MMP-9 (Figure 3.5F).

We next conducted univariate logistic-regression analysis on our data to sort out the role of a number of parameters such as histological grades, molecular subtypes and metastasis on the level of MMP-9 expression. This analysis confirmed the association between the high levels of MMP-9 expression (total scores > 4) with tumors of high histological grade (Grade III) including both HER2-positive and triple-negative breast

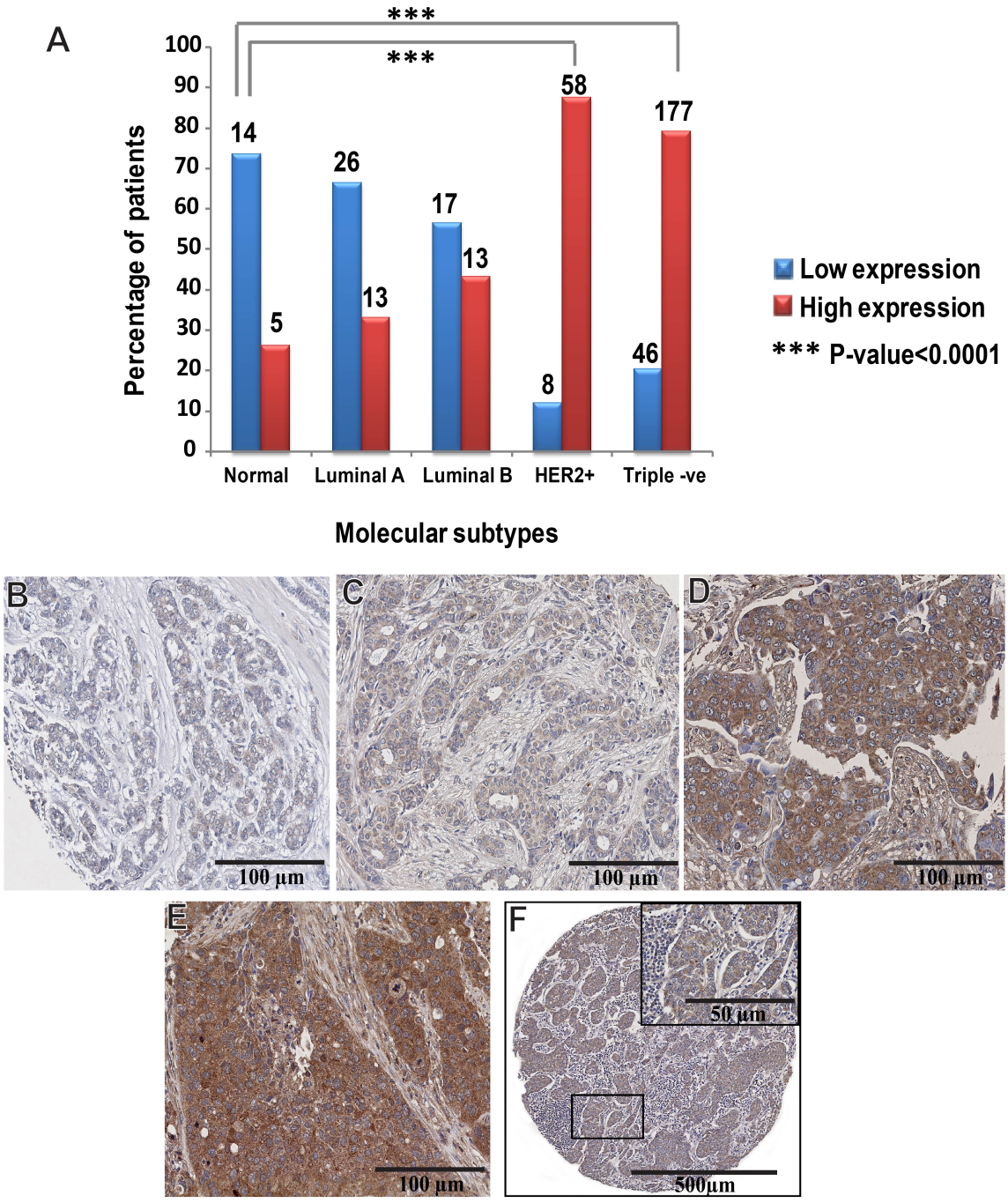

A
***
***
100
90
80
70
60
50
40
30
20
10
0
Percentage of patients
14
5
26
13
17
13
8
58
46
177
Normal
Luminal A
Luminal B
HER2+
Triple -ve
Molecular subtypes
Low expression
High expression
*** P-value<0.0001
B
100 µm
C
100 µm
D
100 µm
E
100 µm
F
50 µm
500µm

Figure 3.5: Overexpression of MMP-9 is associated with triple-negative, HER2-positive breast tumors and nodal metastases.

(A) Histogram showing percentage of breast cancer patients in each molecular subtype category that express low and high level of MMP-9. Both HER2-positive and triple-negative subtypes demonstrate elevated levels of MMP-9 that are significantly different from those observed in normal breast tissue. The number of patients in each group was mentioned over each bar. The overall relationship between MMP-9 scores and molecular subtypes was evaluated using the chi-square test. (**B)** Luminal A and (**C)** Luminal B subtypes showing low level of MMP-9 expression. (**D)** HER2-positive and **(E)** Triple-negative subtypes displaying strong cytoplasmic labeling in cancer cells and surrounding stromal cells. **(F)** Metastatic lymph node demonstrating elevated levels of MMP-9 expression in the cytoplasm of metastatic breast cancer cells. The surrounding lymphocytic and stromal cells did not stain with anti-MMP-9 antibody. Magnification 20X (B-E), 5X (F), 40X inset in Figure 3.5F.

Table 3.1: Univariate analysis of different factors that could affect level of MMP-9 expression.

Parameters	OR	95% CI	*p-value*
Grades			
Grade I	Reference		
Grade II	1.74	0.82-3.73	0.15
Grade III	2.61	1.36-5.08	**< 0.001**
Molecular subtypes			
Luminal A	Reference		
Luminal B	0.51	0.26-0.99	0.05
HER2-positive	8.01	3.85-18.46	**0.001**
Triple-negative	3.90	2.48-6.19	**0.001**
Metastasis (No)	Reference		
Metastasis (Yes)	2.17	1.48-3.23	**0.001**

OR= odds ratio, CI= confidence interval.

cancers (Table 3.1). Hence, we can safely conclude that MMP-9 protein expression *in vivo* strongly supports both *in silico* analyses on microarray dataset as well as data gathered from analysis of breast cancer cell lines.

Overexpression of MMP-9 is associated with a higher incidence of metastases.

We next investigated whether elevated levels of MMP-9 protein expression in carcinoma cells could predict the occurrence of metastases, relapse and poor survival rates. To that end, we reviewed the clinical charts of 200 patients for the period extending from 2000 to 2013. Out of 200 Patients, 121 (60.5%) patients have high MMP-9 expression and 79 (39.5%) patients have low MMP-9 expression. Increased levels of MMP-9 were found to be associated with a higher incidence of metastasis (Figure 3.6). The results were considered significant when the percentage of patients who developed metastases significantly differed in terms of low and high levels of MMP-9 expression. Only lymph node ($p < 0.001$), lymphovascular invasion ($p = 0.007$) and lung metastasis ($p = 0.001$) reached statistical significance when compared to patients with low MMP-9 expression. Additional file 1 indicates the distribution of high and low MMP-9 expression in patients with and without metastases.

Univariate analysis of our data demonstrated the association between elevated levels of MMP-9 expression and the increased likelihood to develop metastasis (OR = 2.17, 95%CI = 1.48-3.23, *p-value* = 0.0001) (Table 3.1). Moreover, to examine which clinical factors could affect the relationship between MMP-9 and metastasis, multivariate logistic-regression analysis was carried out. Triple-negative molecular subtype proved to be the only statistically independent predictor of metastasis (OR = 7.92, 95%CI = 2.90-21.6, *p-value* = 0.0001) (Table 3.2). This suggests that triple-negative breast cancer have a stronger clinical value in predicting metastasis rather than any of the other biological factors examined.

High levels of MMP-9 are associated with a shorter latency to relapse and shorter survival after relapse (SAR)

Likewise, when we looked at the association between MMP-9 and relapse, we found that enhanced expression of MMP-9 was associated with a shorter latency to

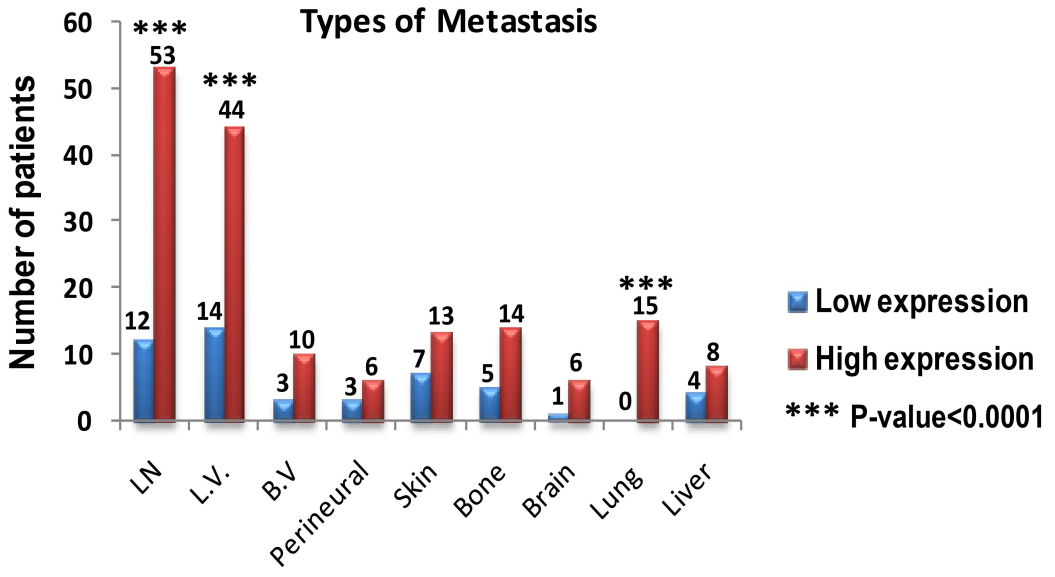


Figure 3.6: Overexpression of MMP-9 is associated with a higher incidence of metastases.

Increased expression of MMP-9 is associated with higher incidence of metastasis. Only lymph node, lymphovascular invasion and lung metastases reached the level of statistical significance when compared to patients with low MMP-9 expression. Chi-square test was realized with Yates' continuity correction and a two-sided Fisher exact test were performed to analyze metastases.

Table 3.2: Multivariate analysis model of MMP-9 that include metastasis, histological subtypes and molecular subtypes.

Parameters	OR	95% CI	*p-value*
Metastasis			
Luminal A	0.97	0.45-2.07	0.93
Luminal B	3.52	0.81-15.27	0.12
HER2-positive	0.77	0.16-3.61	0.79
Triple-negative	7.92	2.90-21.61	**0.001**

OR= odds ratio, CI= confidence interval

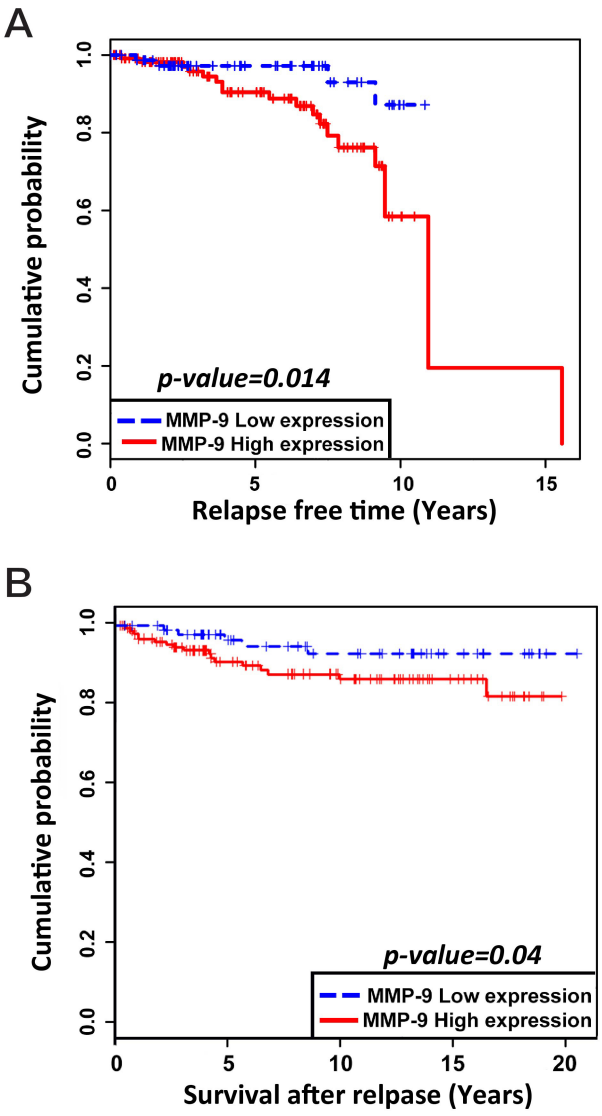


Figure 3.7: Overexpression of MMP-9 is associated with shorter time to relapse and shorter survival after relapse.

(A) High levels of MMP-9 expression are associated with shorter time to relapse (*p=0.014*). **(B)** High levels of MMP-9 expression are associated with shorter survival after relapse (*p=0.04*).

clinical relapse: (Mean time for relapse = 3912 days, n=121) which is statistically significant (p = 0.014). This contrasts with the values observed in patients with low MMP-9 levels of expression (Mean time for relapse = 4957 days, n=79) (Figure 3.7A). However, using a multivariate analysis, histological grades, histological subtypes and molecular subtypes were found to have no impact on relapse in this patient's population.

Finally, the Kaplan-Meier overall survival (OS) curve obtained from the same cohort of patients indicates that increased expression levels of MMP-9 are associated with a shorter OS (Mean OS= 6469 days, n=16) when compared to those tumors expressing low levels of MMP-9 (Mean OS= 6984 days, n=6). However, no significant difference could be identified between OS for patients having high or low levels of MMP-9 expression. Interestingly, univariate analysis shows that patients with higher levels of MMP-9 expression harbor a significant high risk of death after relapse (RR=3.05, p=0.04) (Figure 3.7B). It is also worth mentioning that we could not find any statistically significant correlation between the expression of MMP-9 in the tumor stroma and the occurrence of metastasis or overall survival in the same patients.

Discussion

In the present paper, we provide both indirect and direct evidence that MMP-9 participates to breast cancer progression and impact on clinical outcome. There are many studies reporting the association of elevated levels of MMP-9 with a higher incidence of metastases and poor clinical outcome. We found that high expression of MMP-9 is specifically correlated with high-grade breast cancers that include both triple-negative and HER-2 positive breast cancers.

Previous studies have provided conclusive evidence that MMP-9 is involved in several key processes that contribute to breast cancer development, progression, invasion and metastasis [393, 453, 454]. Here we performed *in silico* analysis of 1210 DNA microarrays of human breast cancer tissues and RNA sequencing data of 51 human breast cancer cell lines to assess *MMP-9* mRNA expression. We found that *MMP-9* mRNA expression in both basal-like and HER2-positive tumors reached significantly higher

levels than those observed in the luminal A category. When the expression of *MMP-9* in breast cancer cell lines is considered, it is worth mentioning that cell lines with a basal-like phenotype and those that are luminal with overexpressed HER2 reached the highest levels of *MMP-9* expression. In contrast, cell lines with luminal phenotype failed to demonstrate elevated levels of *MMP-9*. This strongly suggested to us that MMP-9 expression varied according to cell differentiation and histological grades. Hence, we decided to construct human breast cancer tissue microarrays (TMA) comprising a wide selection of tumors belonging to each category of breast cancer molecular subtypes. Those tumors were classified as triple-negative, HER2-enriched, luminal A and luminal B based on the expression profile of four surrogate markers (ER, PR, HER2, Ki-67) [421]. We also included normal breast tissue to serve as a basis for comparison.

To thoroughly validate the robustness of our IHC assay we first included a number of internal and external controls. Whereas colonic adenocarcinoma strongly expressed MMP-9, two benign breast lesions (fibroadenoma and myofibroblastoma) failed entirely to express MMP-9 under the same conditions. Once the experimental procedures were set up, we performed the IHC assay on TMAs. One important finding was that normal breast tissue displayed either a complete lack of positivity or barely perceptible labeling with the antibody directed against MMP-9. This is consistent with previous observation by others reporting only a weak expression of MMP-9 in normal breast tissue [384, 455]. Indeed, low levels of MMP-9 expression in normal breast tissue are expected since in most tissues MMP-9 is an inducible and not a constitutively expressed gene [456]. Evidently, this sharply contrasts with the high levels of expression of MMP-9 found in the cytoplasm of both HER2-positive and triple-negative breast cancers cells. Hence, our findings support the conclusions of recently published studies indicating a positive correlation between high levels of MMP-9 expression and triple-negative breast cancers [403, 457, 458]. Our results may also explain the findings of La Rocca et al. who showed that high serum levels of MMP-9 are present in HER2 amplified breast cancers [404]. In this context, abnormally elevated levels of MMP-9 can be envisaged as a response to local secretion of inflammatory cytokines and growth factors, such as interleukin 1 (IL-1) and tumor necrosis alpha (TNFα), which may lead to either activation of NF-kB, a well-known inducer of MMP-9 production, or hypomethylation of

its promoter [459]. One cautionary note should be raised though, since high levels of MMP-9 do not necessarily imply high MMP-9 activity as the protein is produced as an inactive pro-enzyme. Moreover, active MMP-9 can be completely neutralized by protease inhibitors such as tissue inhibitors of metalloproteinases (TIMPs) [405]. As for the production site of MMP-9 in breast tumors, our results suggest that carcinoma cells are the main source of MMP-9 given that adjacent stromal cell consistently exhibited a much weaker degree of expression.

Finally, we wanted to correlate clinical outcome characteristics such as onset of metastasis, survival rates and tumor relapse with MMP-9 levels. Our results confirm that overexpression of MMP-9 is tightly correlated with lymphovascular invasion, regional node metastasis, a shorter time to relapse and a reduced SAR. Taken together, our data underscore the role of MMP-9 in promoting breast cancer metastases in lymph node and lungs. This finding is consistent with both *in vitro and in vivo* studies reporting high levels of MMP-9 expression in highly metastatic cell lines [460] and its contribution in metastatic progression [458]. Also, this supports the finding of van 't Veer et al. [156] who demonstrated in a DNA microarray study that MMP-9 is significantly upregulated in poor prognosis signature of breast cancer. Although we have not directly addressed the question on how MMP-9 fosters invasion and nodal metastasis, there are numerous conceivable explanations that can be put forth such as alteration of basal membrane components, diminished cell-to-cell adhesion, release of ECM-bound growth factors and chemotactic molecules, stimulation of angiogenesis and induction of the epithelial-mesenchymal transition (EMT) [387, 392, 461-463].

At any rate, our findings clearly emphasized the clinical potential of MMP-9 as a prognostic biomarker in breast cancer. This is in agreement with Wu et al. [464] who suggested the potential role of MMP-9 as a biomarker for breast cancer progression. Interestingly, the first fully commercialized and FDA approved microarray-based multigene assay for breast cancer, MammaPrint®, does include MMP-9 among its 70 panel genes [272, 465]. Given on the one hand the overwhelming interest in developing prognostic and predictive breast cancer assays and, on the other, the recognition that so called “wound-healing” or “invasion” gene signatures are important to predict tumor

relapse and benefit to chemotherapy, one might consider including MMP-9 alone or in combination with other genes in the development of other multigene multiplex assays.

Conclusion

In summary, our results indicate that overexpression of MMP-9 is closely associated with breast cancers of high histological grade including triple-negative and HER2-positive molecular subtypes. Increased levels of expression of MMP-9 are also correlated with the onset of nodal metastases, a reduced time interval to relapse and a shorter SAR. Taken together, our findings suggest that the differential expression of MMP-9 contributes to breast cancer heterogeneity and is a key characteristic of the "molecular signature" of subsets of breast cancer. In our opinion, MMP-9 expression could help segregate subsets of aggressive breast cancer into clinically meaningful subtypes. Lastly, our results suggest that MMP-9 is a valuable gene/protein candidate to be considered in the development of a multi-gene panel or multiplex proteomic assay to predict clinical outcome.

Additional materials

Additional file 1: Number of patients with or without metastasis associated with either high or low MMP-9 expression.

	Metastasis	High Expression	Low Expression
Lymph node (L.N.)	Yes	53	12
	No	68	67
Lympho-vascular (L.V.)	Yes	44	14
	No	77	65
Blood vessels (B.V.)	Yes	10	3
	No	111	76
Perineural	Yes	6	3
	No	115	76
Skin	Yes	13	7
	No	108	72
Bone	Yes	14	5
	No	107	74
Brain	Yes	6	1
	No	115	78
Lung	Yes	15	0
	No	106	79
Liver	Yes	8	4
	No	113	75

CHAPTER IV

MCM2: an alternative to Ki-67 for measuring breast cancer cell proliferation

Einas M Yousef[1], David Laperriere[1], Muhammad R Tahir[2], Sylvie Mader[1&3], Louis A Gaboury[1&4§]

[1]Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, Canada, [2]Le centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Canada, [3]Department of Biochemistry, Faculty of Medicine, Université de Montréal, Montréal, Canada, [4]Department of Pathology and Cell Biology, Faculty of Medicine, Université de Montréal, Montréal, Canada

§Corresponding author

Submitted to Breast cancer research Journal in April 2015 and it is under review.

Authors' contributions

EY, SM and LG are responsible for the study design.

EY performed the experiments and collected the data.

EY and DL performed the *in silico* analysis.

EY and LG scored the IHC reactions.

EY prepared all the diagrams and pictures.

EY and MT carried out the statistical analysis.

EY, SM and LG participated in the data analysis and interpretation.

All authors drafted, read and approved the final manuscript.

Abstract

Background

Breast cancer is a heterogeneous disease comprising a diversity of tumor subtypes that manifest themselves with a wide variety of clinical, pathological and molecular features. One important subset, luminal breast cancers (ER+/HER2-), comprises two clinically distinct subtypes: luminal A and luminal B each of them endowed with its own genetic program of differentiation and proliferation. Luminal B tumors tend to be less differentiated and exhibit a higher rate of proliferation than luminal A breast cancers. There is currently a need for a clinically robust and validated immunohistochemical assay to help distinguish between luminal A and B breast cancer. MCM2 is a family member of the mini-chromosome maintenance protein complex whose role in DNA replication and cell proliferation is firmly established. It has been reported to be a promising proliferative marker in many types of cancer including thyroid, rectal and breast cancers. As MCM2 appears to be an attractive alternative to Ki-67, we sought to study the expression of MCM2 and Ki-67 as they relate to different histological grades and molecular subtypes of breast cancer focusing primarily on ER-positive tumors.

Methods

MCM2 and *Ki-67* mRNA expression were first studied using *in silico* analysis of available DNA microarray and RNA sequencing data of human breast cancer tissues. We next used immunohistochemistry to evaluate protein expression of MCM2, Ki-67 and a wide array of breast cancer biomarkers on tissue microarrays constructed from a cohort of 266 breast cancer patients

Results

Both MCM2 and Ki-67 are highly expressed in breast cancer of high histological grades that comprise clinically aggressive tumors such as triple-negative, HER2-positive and luminal B subtypes of breast cancer. MCM2 expression was also found to express at higher levels than Ki-67 ($p < 0.0001$) in normal breast tissues, and in breast cancers. Two

distinct subgroups among ER+/HER2- breast cancer could be easily identified using a MCM2 threshold of 40%.

Conclusion

MCM2 expression can help in distinction between two clinically important subsets of luminal breast cancer whose treatment and clinical outcomes are likely to diverge. An independent validation cohort is needed to confirm the clinical utility of MCM2.

Keywords: MCM2, Ki-67, human breast cancers, *in silico* analysis, tissue microarray (TMA)

Competing interests

The authors declare that they have no competing interests.

Introduction

According to the last annual report of the American Cancer Society, it is estimated that 232,670 new cases of invasive breast cancer will be diagnosed among women in the United States in 2014 resulting in an estimated 40,000 breast cancer deaths. Only lung cancer exacts a larger toll from women in terms of mortality [19]. Breast cancer is a heterogeneous disease that manifests itself with an embarrassing variety of morphological characteristics, a complex array of clinical and pathological features, contrasting molecular subtypes, as well as variable responses to treatment [421, 466].

Molecular subtypes of breast cancer, primarily identified by gene expression profiling, harbor distinct gene expression patterns that translate into characteristic disease phenotypes with significant differences in outcomes [36, 467]. Molecular classification can, to a certain extent, be clinically approximated using IHC expression profiles of four surrogate markers: ER, PR, HER2 and Ki-67 [23, 44, 50]. Molecular subtypes include estrogen receptor positive (ER+) breast cancer (luminal A and luminal B) and estrogen receptor negative (ER-) breast cancer (HER2-positive, basal-like and a so-called normal breast-like breast cancer subtype) [468]. Luminal A and luminal B breast cancers are significantly distinct at the molecular level as luminal A breast cancer express the highest level of ER-related and the lowest level of proliferation-related genes, while luminal B breast cancer show the opposite pattern of gene expression [41, 44, 45].

The proliferation signature of tumor cells, comprised of cell-cycle regulated genes whose expression correlates with rapid cell proliferation [469, 470], has been detected in different types of cancer such as breast, prostate, liver and lung [43, 471-473]. There is now growing evidence that the proliferative capacity of breast cancer impacts on the prognosis, clinical behavior and aggressiveness of the tumor suggesting that accurate measurement of cell proliferation may help selecting a more appropriate treatment [280, 474, 475]. Indeed, proliferation is considered as the highest weight component in the Oncotype DX recurrence score (RS) [476]. While DNA microarray is an ideal way to measure multi-gene expression proliferation signature in a quantitative and automated manner, the need for a histological-based proliferation test is crucial for pathologist routine clinical assessment of breast cancer [469, 477]. Currently, mitotic index and

immunohistochemistry of proliferation-associated antigen such as Ki-67, cyclins, minichromosome maintenance (MCM) proteins and topoisomerase IIα are the most common practical histological methods to assess proliferation [474, 478].

Ki-67 is one of the most important cell proliferation related genes and, as such, it has been widely investigated as a proliferative marker in breast cancer with the hope of distinguishing luminal A from luminal B breast cancer [50, 479, 480]. However, Ki-67 assessment in breast cancer has failed to impose itself as an accepted proliferative biomarker because of a lack of reproducibility and the difficulty to establish an appropriate cutoff (i.e. 10%, 14% and 20%) [45, 50, 282, 327-329]. Indeed, neither the ASCO nor the NCCN has included Ki-67 in its recommended list of routine breast cancer biomarkers [281, 330]. Furthermore, the IMPAKT 2012 working group pointed out that there was not enough evidence to support using a Ki-67 labeling index of 14% to identify clinically relevant subtypes of breast cancer [112]. In this context, it is undeniable that identification of a robust breast cancer proliferation marker would be of the utmost interest. MCM2, whose role in initiating DNA replication is now firmly established [283], appears to be an appealing alternative to Ki-67.

MCM2 belongs to the mini-chromosome maintenance (MCM) proteins complex or pre-replicative complex, which consists of six highly conserved proteins (MCM2-MCM7) collectively interacting to bring about initiation of DNA replication [334]. This complex is considered to be the converging point that connects the growth signaling pathways with the initiation of DNA replication [481]. MCMs are expressed in actively proliferating cells and non-cycling cells with proliferative potential [482]. Recently, it has been recognized that cancers arising in different anatomic sites such as stomach, colon and skin are associated with MCM2 overexpression [483-485]. Furthermore, recent reports are now ready to support the notion that MCM2 is a promising proliferative marker in many other types of cancer including thyroid, rectal and breast cancers [348, 349, 351].

In this study, we reasserted the expression of MCM2 and Ki-67 as they relate to breast cancers of different histological grades and molecular subtypes focusing primarily on ER-positive tumors. To achieve that aim, we first studied *MCM2* and *Ki-67* mRNA expression using *in silico* analysis on available DNA microarray and RNA sequencing

data of human breast cancer tissues. We next used immunohistochemical (IHC) staining to evaluate expression of MCM2, Ki-67 and a wide array of breast cancer biomarkers on tissue microarrays (TMAs) constructed from a cohort of 266 breast cancer patients. Our results indicate that both MCM2 and Ki-67 are highly expressed in breast cancer of high histological grades that comprise clinically aggressive tumors such as triple-negative, HER2-positive and luminal B subtypes of breast cancer. MCM2 expression was also found to express at higher levels than Ki-67 ($p < 0.0001$) in normal breast tissues, and in breast cancers of different histological grades and molecular subtypes. Notably, two distinct subgroups among hormonally responsive luminal breast cancer (ER+/HER2-) could be easily identified using a MCM2 threshold of 40% which is easier and convenient to use than the unyielding 14% index of Ki-67. Importantly, a bimodal frequency distribution of MCM2 scores could be conclusively demonstrated among luminal breast cancers. In brief, our data strongly support the notion that MCM2 expression can help teasing out two clinically important subsets of luminal breast cancer whose treatment and clinical outcomes are likely to diverge. Evidently, an independent validation cohort is needed to confirm the clinical applicability of MCM2 to tailor breast cancer treatment.

Material and methods

In silico Analysis

The web application bc-GenExMiner [415] was used for correlation analysis of *MCM2* and *MKI67* gene expression on a dataset comprising over 3,063 microarrays. However, using the criteria selected for this application, only 1260 microarrays could be positively ascribed to one of the five different molecular subtypes of breast cancer. Study into the differential expression of mRNA level of both *MCM2* and *MKI67* between different molecular subtypes of breast cancer was carried out on the same cohort of patients. Gene expression values were determined beforehand and split in order to define three equal groups (low, intermediate and high expression). This means that “high

expression" is the 1/3 of the patients with highest expression of *MCM2* or *MKI67* and "low expression" is the lower 1/3 of the patients.

To validate results obtained from bc-GenExMiner application, MiSTIC dataset (RNAseq data of 756 breast cancer patients) and The University of California Santa Cruz (UCSC) Cancer Genomics Browser (Gene expression array of 597 patients (AgilentG4502A_07_3 array)) were also used [486]. The heat map (Additional file 1) produced from the USCS cancer genomic browser displays the expression of different genes using distinct sets of colors, red representing data values > 0, green values < 0 and black value = 0. The "aov" and "TukeyHSD" functions were carried out to compare the mRNA levels within each breast cancer molecular subtype. The ANOVA was applied to check for an overall difference of expression levels between each molecular subtype. The Tukey multiple comparisons of means were used to test for a significant difference between two subtypes (e.g. luminal A vs. luminal B). For both tests, a *p-value* < 0.05 was considered significant. Pearson correlation coefficient between set of genes was calculated with the "cor.test" function of the R language and environment for statistical computing (http://www.R-project.org/).

Patients and Tissue Samples

A cohort of 266 female breast cancer patients comprising tumors of different histological grades was selected for the present study. Formalin-fixed, paraffin-embedded (FFPE) samples containing tumor tissues were collected after surgery (lumpectomy or mastectomy) (Table 4.1). Tumor grades were confirmed using the Modified Scarff-Bloom-Richardson-Elston-Ellis grading system (SBR-EE) [444]. A complete set of follow-up data including the onset of relapse and death was available for review. We also added 21 normal breast tissues from healthy women undergoing plastic surgery to serve as internal controls. Normal human brain was included as negative control for MCM2 [403]. In addition, a number of extraneous tissues such as colon, thyroid and placenta were included in each TMA. All samples were obtained from *Centre Hospitalier de l'Université de Montréal* (CHUM) after granting the approval of the research ethical committee (SL 05.019). Since all donor blocks remained anonymous no individual patient consent was required.

Table 4.1: **Clinico-pathological data of tissues used in TMAs**

Variables	No. of cores	%
Organs used in TMA	469	
Colon	3	0.64
Lymph node	26	5.54
Mammary gland • Normal breast tissue • In situ carcinoma • Benign breast tumors • Invasive breast cancer	421 21 14 6 380	89.77
Placenta	1	0.21
Rectum	12	2.56
Thyroid	6	1.28
Grade	355	
I	32	9.01
II	63	17.75
III	260	37.24
Molecular subtypes	364	
Luminal A	108	29.67
Luminal B	48	13.19
HER2-positive	64	17.58
Triple-negative	144	39.56

Tissue Microarray (TMA)

Sections (4 μm) from each paraffin donor block were stained with hematoxylin and eosin (H&E) and examined by two independent pathologists. Core punches, 1mm in diameter, were plucked from representative areas contained within each FFPE tumor blocks. Each core was realigned in duplicate or triplicate into recipient blocks according to the intended design of the map using a Manual Tissue Arrayer I (Beecher Instruments). Blocks were next inverted and incubated overnight in the oven over a glass slide. TMA blocks were allowed to cool until they could easily detach from the glass slide. Tissue sections from each TMA were prepared and one slide from each block was stained with H&E to review the diagnoses and histological grades on all tissue samples. Additional representative sections from each block were submitted to IHC staining.

Immunohistochemistry

IHC assays were performed on FFPE tissues obtained from each TMA. These assays were carried out according to manufacturer recommendations on an automated immunostainer (Discovery XT system, Ventana Medical Systems, Tucson, AZ). IHC analysis of MCM2 (monoclonal; #12079, dilution 1/200, citrate buffer (sCC1), Cell signaling) was carried out. Specificity of the anti-MCM2 monoclonal antibody was confirmed using normal colonic mucosa [487] and normal cerebral cortex as positive and negative controls, respectively (Additional file 2), based on the expression data in the Human Protein Atlas database [14]. In addition, IHC analysis of estrogen receptor (ER; monoclonal, clone SP1, RTU, sCC1, Ventana Medical Systems), progestrone receptor (PR; monoclonal, clone 1E2, RTU, sCC1, Ventana Medical Systems), HER2 (monoclonal, clone 4B5, RTU, sCC1, Ventana Medical Systems), Ki-67 (monoclonal, clone SP6, dilution 1/100, pretreated sCC1, BioCare medical) was carried out as previously described [488].

Scoring of stained slides

The scoring systems used for each antibody are listed in additional file 3. The expression of MCM2 and Ki-67 in breast epithelium was studied by calculating the percentage of positively stained nuclei [50, 344]. Two different approaches of scoring were used, a visual method using conventional light microscopy and computer-assisted

automated scoring method using Visiomorph©, Tissuemorph© Digital Pathology (DP) softwares. Visual scoring of MCM2 and Ki-67 staining reactions is labor-intensive, time consuming and is subjected to intra-observer and inter-observer variability. To increase reproducibility in the visual scoring system, evaluation was performed by two independent observers. Automated MCM2 and Ki-67 scorings were also carried out resulting in a much more rapid and accurate readout. Visiomorph© DP has the distinct advantage of leaving out stromal cells from the analysis retaining only cancer cells in the Region Of Interest (ROI). As for Tissuemorph© DP, it allows accurate counting of the positive and negative nuclei in the ROI. ER, PR, HER2 and Ki-67 were also used as surrogate markers to sub-classify breast cancers into different molecular subtypes as listed in additional file 4 [112, 421].

Statistical Analyses

All statistical analyses were carried out using XLSTAT (http://www.xlstat.com/en/) and different packages of the R language (http://www.R-project.org/). The allocation of MCM2 and Ki-67 positive cases among different histological grades and molecular subtypes is depicted using boxplots and scatter plots for both visual and automated cell counts. The overall relationship between MCM2 and Ki-67 scores and both histological grades and molecular subtypes was evaluated using the chi-square test and Mann-Whitney test / Two-tailed test. Correlation analysis for IHC expression levels was carried out using the Spearman's rho correlation coefficient for visual scores and Pearson correlation coefficient for automated score. Concordance of automated and visual scoring was also achieved using Intra-class Correlation Coefficient (ICC) [489].

The Receiver-Operating Characteristic (ROC) curve was used to detect the optimal cutoff point, which simultaneously reached maximum sensitivity and specificity values for MCM2 score. This could be achieved only for the automated MCM2 scores. Using this cutoff point, continuous variables could then be treated as dichotomous variables (low and high MCM2 expression) [490]. Distribution of MCM2 and Ki-67 in luminal breast cancer was displayed using histograms. Kaplan-Meier plot was drawn to show the overall survival and survival after relapse for low and high levels of MCM2

expression. Cox regression was used for multivariate survival analysis. Statistical significance was considered with a *p-value* less than 0.05.

Results

MCM2 is highly expressed at the mRNA level in basal-like, luminal B and HER2-positive but not luminal A breast cancer.

The web application bc-GenExMiner was used to compare the mRNA levels within each breast cancer molecular subtype on a dataset comprising 1260 microarrays [415]. As shown in Figures 4.1A, 1B & additional file 1, basal-like, luminal B and HER2-positive breast cancers show higher expression of *MCM2* and *MKI67* mRNA levels in comparison with both luminal A and normal-like breast cancers ($P < 0.0001$). Notably, there is no overlap between individual boxes in the boxplot when luminal A were compared to luminal B breast cancers (Figure 4.1A & 1B). The percentage of patients with low, intermediate and high levels of *MCM2* and *MKI67* expression in each molecular subtype of breast cancer is depicted in figure 4.1C. In this microarray data set, 74% of basal-like (total no. = 388), 62% of luminal B (total no. = 116) and 49% of HER2-positive (total no. = 103) breast cancer patients show high level of *MCM2* mRNA expression. In contrast, only 7% of luminal A (total no. = 443) and 5% of normal-like (total no. = 210) breast cancer show high level of *MCM2* mRNA expression. Similar figures were observed in the different molecular subtypes of breast cancer using *MKI67* mRNA expression data (Figure 4.1C).

MCM2 was next correlated with *MKI67* and *ESR1* in breast cancers based on RNA-sequencing data derived from The Cancer Genome Atlas (TCGA). As predicted, a strong positive correlation could be detected between *MCM2* and *MKI67* ($r = 0.73$), whereas a negative correlation was observed between *ESR1* and both *MCM2* and *MKI67* ($r = -0.33$, $r = -0.38$ respectively) (Figure 4.1D). This close similarity between the *in silico* profiles of *MCM2* and *MKI67* prompted us to pursue assessment of MCM2 protein expression in breast carcinoma tissue microarrays along with cognate normal breast tissue.

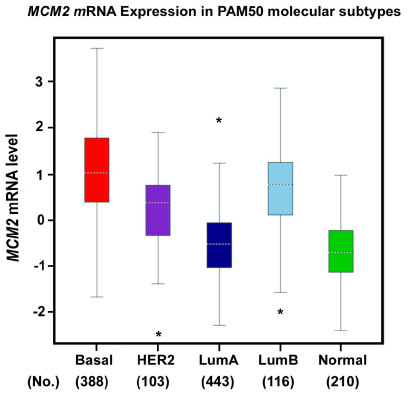
High concordance between automated and visual scores

After confirming MCM2 immune reactivity, sensitivity and specificity, we carried out high-throughput IHC reactions on TMAs comprising normal, benign and malignant breast tissues. Two different approaches were used to calculate the percentage of positive nuclei of MCM2 and Ki-67 in the stained sections, a visual scoring and an automated quantifying method. Intraclass Correlation Coefficient (ICC) was calculated to compare visual and automated scores and to evaluate their relative performance [489]. ICC was obtained for the 365 pairs of automated and visual scores of MCM2 and Ki-67. Our results demonstrated that ICCs for MCM2 automated versus visual scoring were 0.94 (95% CI= 0.93-0.95, $p < 0.0001$) for single measures and 0.97 (95% CI= 0.96-0.97, $p < 0.0001$) for average measures. ICCs for Ki-67 automated versus visual scoring were 0.90 (95% CI= 0.87-0.91) for single measures and 0.94 (95% CI= 0.93-0.95) for average measures. Therefore, there is an excellent agreement (ICC > 0.90) between automated and visual scorings for both MCM2 and Ki-67 [489]. Taken together, our results suggest that both methods can be used to reliably assess MCM2 and Ki-67 expressions in breast cancer.

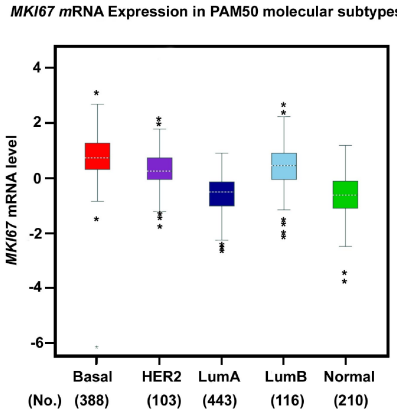
A 40% index cutoff point can distinguish between two distinct subgroups of low and high expression of MCM2 in breast cancer

A ROC curve was used to set the optimal cutoff point based on the continuous automated MCM2 data set. When the accuracy and the sum of sensitivity and specificity were taken into account, the optimal cutoff point corresponded to a value of 40%. Using this cutoff value, the sensitivity of MCM2 index was found to be 69% and specificity 91%. We next evaluated the accuracy of the test by measuring the Area Under the Curve (AUC) as it provides clear indication on how well the ROC curve separates out the group being tested into two distinct subgroups. When our scores were considered, the approximate AUC was found to be 0.81 ($p < 0.0001$, 95% CI = 0.221–0.391). This implies that a 40% cutoff point can be considered a “good” discriminating value in

A



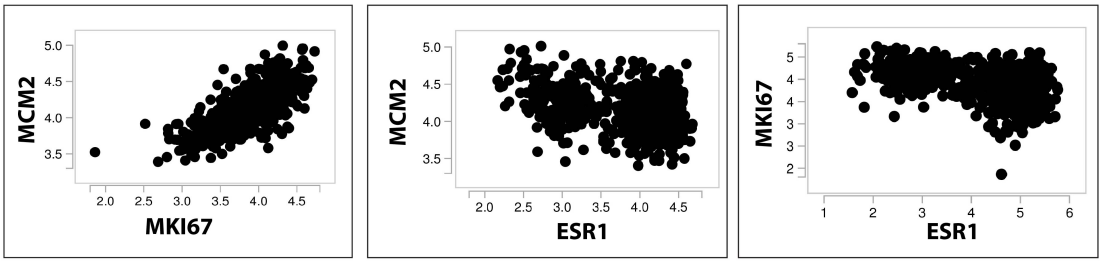
B



C

Molecular subtype	Basal-Like			HER2+			Luminal A			Luminal B			Normal breast -Like		
No.	388			103			443			116			210		
	Low	Inter.	High	Low	Inter.	High	Low	Inter.	High	Low	Inter.	High	Low	Inter.	High
MCM2 Expression	5%	20%	74%	23%	28%	49%	55%	37%	7%	12%	26%	62%	64%	30%	5%
MKI67 Expression	5%	20%	75%	16%	37%	48%	61%	35%	4%	16%	23%	60%	62%	32%	6%

D



r(log10(1000*RPKM+1))= 0.73

r(log10(1000*RPKM+1))= -0.33

r(log10(1000*RPKM+1))= -0.38

Figure 4.1: *In silico* analysis of *MCM2* and *MKI67* mRNA expression in PAM50 molecular subtypes of breast cancer

A) & B) Boxplots depicting higher levels of *MCM2* and *MKI67* mRNA levels in basal-like, luminal B and HER2-positive breast cancers in comparison to luminal A and normal-like breast cancers (p-value <0.0001). Notably, there is no overlap between individual boxes when luminal A were compared to luminal B breast cancers. **C)** Representation from bc-GenExMiner database v3.0.A showing the number of patients with low, intermediate and high *MCM2* and *MKI67* expression in each molecular subtype of breast cancer. Elevated expression of *MCM2* and *MKI67* is present in basal-like, luminal B and HER2-positive breast cancer. **D)** Correlation diagrams using MiSTIC visualization tool to correlate *MCM2* with *MKI67* and *ESR1* in breast cancers based on RNA-sequencing data derived from The Cancer Genome Atlas (TCGA). A strong positive correlation could be detected between *MCM2* and *MKI67* (r = 0.73), whereas a negative correlation is observed between *ESR1* and both *MCM2* and *MKI67* (r = -0.33, r = *-0.38* respectively).

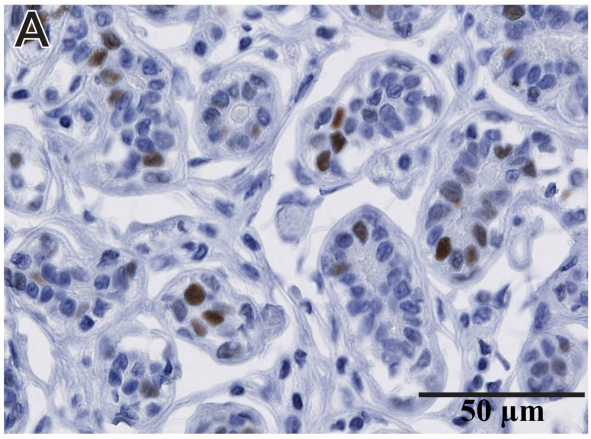
separating low and high MCM2 expression levels [491]. In other words, tumors with scores ranging from 0-40% could be considered to have low MCM2 expression while those with scores exceeding 40% were considered to have high MCM2 expression. As for Ki-67, we maintained the canonical 14% threshold used in the literature [44, 50]. Of note, Cheang et al. used a similar approach to determine the cutoff of Ki-67 to distinguish luminal A from luminal B breast cancer [50].

Two distinct patterns of MCM2 expression were detected in normal breast tissues

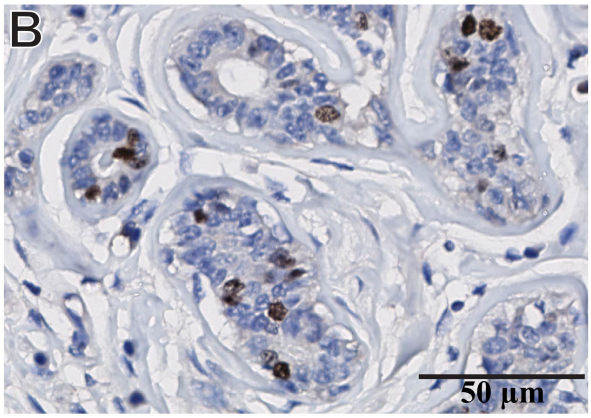
The majority (62%) of normal breast tissue samples exhibit MCM2 nuclear labeling in only a few scattered luminal cells lining the terminal duct lobular units (Figure 4.2A). This pattern of expression was also found in 100% of normal breast tissues using Ki-67 (Figure 4.2B). However, as many as 36% of normal breast tissues demonstrated a strong MCM2 nuclear labeling in most if not all of the luminal cells lining the normal breast ducts (Figure 4.2C). Of note, there was no MCM2 or Ki-67 labeling neither in the myoepithelial cell layer nor in the surrounding stromal cells. Also, we never observed cytoplasmic or membranous staining in any of the labeled cells. Although a positive correlation could be established between MCM2 and Ki-67 scores in normal breast tissue (r = 0.45, p = 0.04), MCM2 was expressed at higher levels in normal breast tissue compared to Ki-67 (Mean difference between two scores = 40%, 95% CI = 22.21-56.94, $P < 0.0001$).

Higher levels of expression of MCM2 are associated with breast cancers of high histological grades.

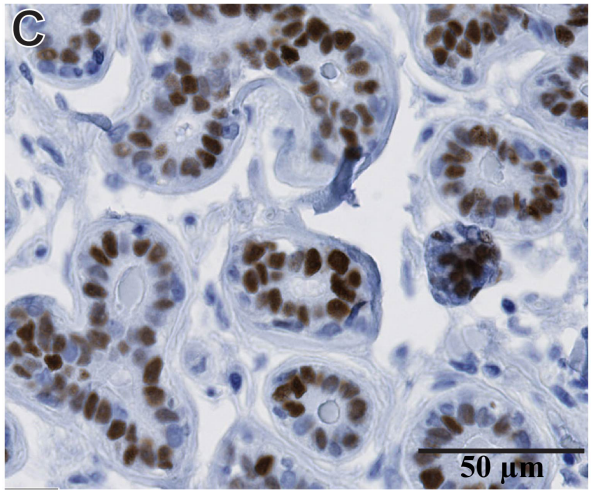
The expression of MCM2 and Ki-67 was then examined in breast cancers of various histological grades. Based on the previously defined 40% cutoff, our results demonstrated that 97% of grade I breast cancer expressed low levels of MCM2. On the other hand, high levels of MCM2 expression were detected in 29% of grade II ($p < 0.0001$) and 74% of grade III breast cancer ($p < 0.0001$) which is statistically significant when compared to grade I breast cancer (Figure 4.3A-D). We also found significantly higher levels of expression of Ki-67 in grade II ($p < 0.014$) and grade III breast cancers ($p < 0.0001$) in comparison to grade I tumors (Figure 4.3E).



MCM2 expression in 62% of normal breast tissue



Ki-67 expression in 100% of normal breast tissue



MCM2 expression in 36% of normal breast tissue

Figure 4.2: Patterns of MCM2 and Ki-67 expression in normal breast tissue

A) The majority (62%) of normal breast tissue samples displays nuclear labeling in only a few scattered luminal cells lining the terminal duct lobular units. **B)** All normal breast tissues labeled with Ki-67 have shown nuclear labeling in only a few scattered luminal cells lining the terminal duct lobular units. **C)** In contrast, 36% of normal breast tissues demonstrated a strong MCM2 nuclear labeling in most if not all of the luminal cells lining the normal breast ducts. There was neither MCM2 nor Ki-67 labeling either in the myoepithelial cell layer or in the surrounding stromal cells. Of note, we never observed cytoplasmic or membranous staining in any of the labeled cells. Magnification 20X

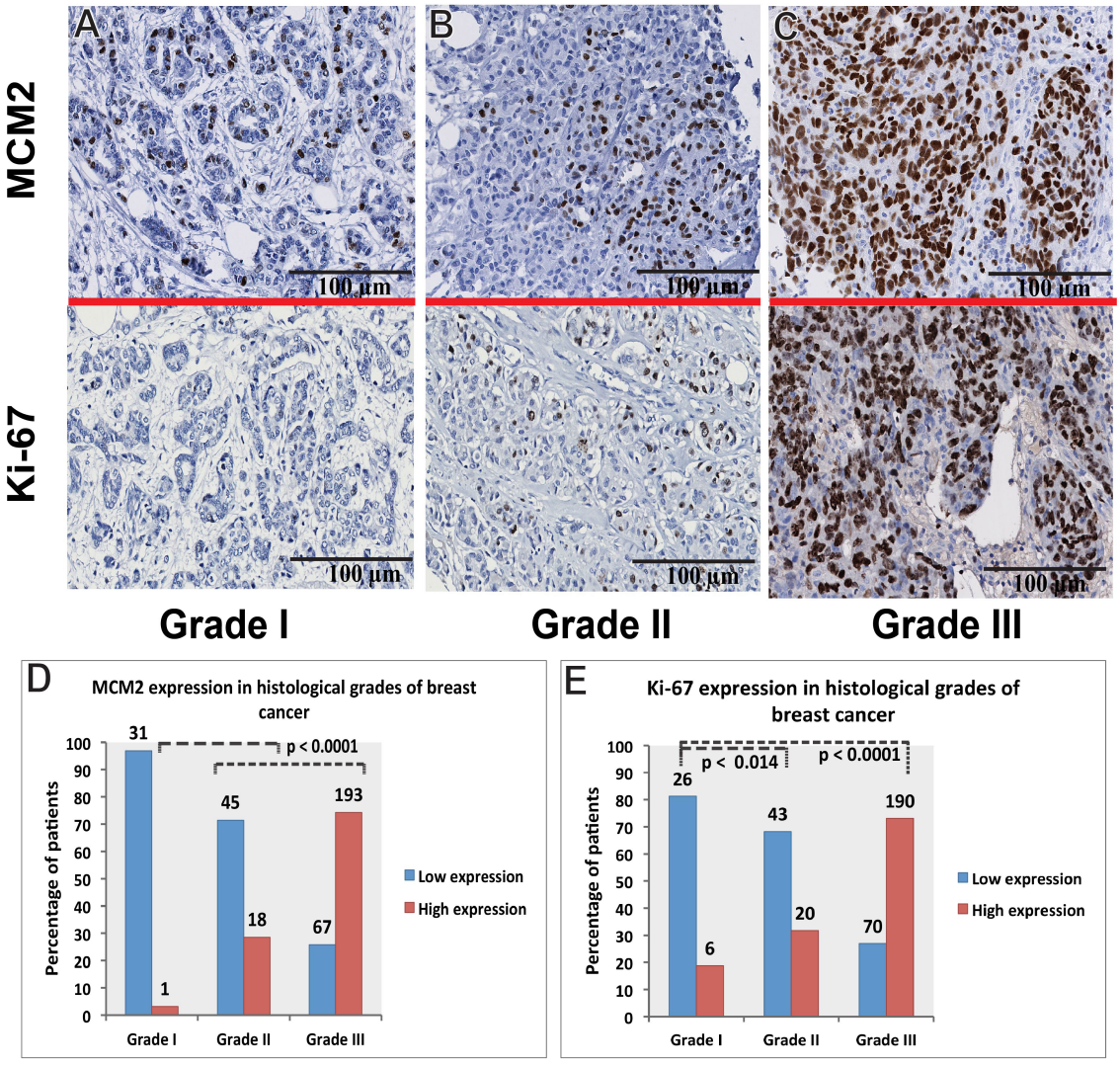

MCM2
Ki-67
A
B
C
100 µm
Grade I
Grade II
Grade III
D
MCM2 expression in histological grades of breast cancer
Percentage of patients
p < 0.0001
31
1
45
18
67
193
Low expression
High expression
Grade I
Grade II
Grade III
E
Ki-67 expression in histological grades of breast cancer
p < 0.014
p < 0.0001
26
6
43
20
70
190
Low expression
High expression
Grade I
Grade II
Grade III

Figure 4.3: Overexpression of MCM2 is associated with high histological grade of breast cancer.

A) Grade I, **B)** Grade II, **C)** Grade III breast cancer stained with MCM2 (upper panels), Ki-67 (lower panels). Overexpression of MCM2 and Ki-67 is associated with breast cancer of higher histological grade. Levels of MCM2 expression characteristically exceeded that of Ki-67 in different grades of breast cancers (pictures were taken from the same area of adjacent slides). D) Histograms displaying the percentage and number of patients with low and high MCM2 expression in breast cancer of different histological grade. Most of grade I breast cancer (97%) expressed low levels of MCM2. On the other hand, high levels of MCM2 expression were detected in 29% of grade II ($p < 0.0001$) and 74% of grade III breast cancer ($p < 0.0001$) which is statistically significant when compared to grade I breast cancer. E) Histograms displaying the percentage and number of patients with low and high Ki-67 expression in different histological grade of breast cancer. Significant higher levels of Ki-67 expression in grade II ($p < 0.014$) and grade III breast cancers ($p < 0.0001$) in comparison to grade I tumors was also detected.

As far as the cellular localization is concerned, MCM2 and Ki-67 reactivity was restricted to the nuclei of cancer cells. Neither of the two markers was detected in the adjacent stromal cells. Again, levels of MCM2 expression in breast cancers exceeded that of Ki-67 (Mean difference between two scores = 25%, 95% CI = 20.12-29.54, *P-value* < 0.0001) (Figure 4.3 A-C). As expected, we could find a positive correlation between MCM2 and Ki-67 in breast cancer of different histological grades (Grade I $r = 0.69$, Grade II $r = 0.73$, Grade III $r = 0.91$, $p < 0.0001$). We next addressed the question of Ki-67 cell labeling variability and heterogeneity in tumors since this issue has been raised previously for different types of cancers [492, 493]. Among the 380 representative breast cancer cores that we analyzed for Ki-67 and MCM2, 20 cores (5.3%) show marked intra-tumoral heterogeneity using Ki-67. Notably, MCM2 labeling in the same regions failed to show any significant degree of heterogeneity (Figure 4.4).

Increased expression of MCM2 is associated with luminal B, triple-negative, and HER2-positive breast cancer

To validate the results of the *in silico* analysis in breast tumors, we studied the expression of MCM2 and Ki-67 at the protein level in breast cancer of different molecular subtypes and assessed their respective cellular and subcellular localization. Molecular subtypes were operationally defined using the following four surrogate markers: ER, PR, HER2 and Ki-67 as shown in additional file 4 [112]. As shown in table 4.2, 95% of luminal A breast cancer expressed low levels of MCM2 with only 5% showing high expression of MCM2. In contrast, high levels of MCM2 expression were detected in luminal B (79%, $p < 0.0001$), HER2-positive (81%, $p < 0.0001$) and triple-negative (81%, $p < 0.0001$). As far as MCM2 is concerned, all subtypes display significant differences when compared to luminal A breast cancer (Figure 4.5A). Likewise, significant differences could be detected between Ki-67 expression in luminal B, HER2-positive and triple negative breast cancer ($p < 0.0001$) when compared to luminal A subtype (Figure 4.5B). As mentioned earlier, levels of MCM2 expression are significantly and consistently higher than those of Ki-67 for each molecular subtype. The mean difference between the two scores in luminal A subtype is 13.10% (95% CI = 0.01-16.02, *P-value* < 0.0001), luminal B subtype is 18.98% (95% CI = 0.008-30.11, *P-value*

= 0.001), HER2-positive breast cancer is 24.99% (95% CI = 0.015-34.67, *P-value* < 0.0001) and triple-negative breast cancer is 18.67 % (95% CI = 0.01-25.65, *P-value* < 0.0001) (Figure 4.5C-F).

In agreement with the *in silico* correlation analyses, we found a negative correlation between MCM2 and ER (r = -0.54, $p < 0.0001$) and MCM2 and PR (r = -0.49, $p < 0.0001$) in breast cancer at the level of protein expression. In contrast, we could establish a positive correlation between MCM2 and Ki-67 in different molecular subtypes of breast cancer (luminal A $r = 0.55$ ($p < 0.0001$), luminal B $r = 0.46$ ($p = 0.001$), HER2-positive $r = 0.65$ ($p < 0.0001$), triple-negative $r = 0.62$ ($p < 0.0001$)).

MCM2 scores in luminal breast cancer follow a bimodal frequency distribution.

Luminal A (108 patients) and luminal B (48 patients) breast cancer were defined by positive expression of ER and/or PR, lack of HER2 expression and 14% labeling index of Ki-67. As shown in figure 4.6, these subtypes are depicted in black for luminal A and in red color for luminal B tumors. The frequency distribution of MCM2 scores for all luminal breast cancer (156 patients) was carried out. It is worth mentioning that we could find a bimodal distribution of MCM2 scores in luminal breast cancer suggesting that MCM2 can separate out two distinct subgroups among hormonally responsive luminal breast cancers (Figure 4.6A). Indeed, among the 156 samples of luminal breast cancer, 70% showed scores ranging from 0 to 38.4% that represent luminal breast cancer with low proliferation. On the other hand, approximately 30% of patients showed MCM2 scores ranging from 38.5-100% defining a class of luminal breast cancer endowed with high proliferation. This distribution pattern readily emphasizes the usefulness of a 40% threshold, which is very close to 38.4%, as a cutoff point to distinguish between MCM2 high and low expression. When the same type of analysis was applied to Ki-67 scores using the same set of 156 luminal breast cancer tissue cores, we failed to observe a similar bimodal distribution in luminal breast cancers (Figure 4.6B).

As shown in figure 4.6A, one can appreciate that some of the luminal A breast cancer, as defined by 14% index of Ki-67, are present in the area of MCM2 high expression while some of luminal B breast cancer, are found in the area of MCM2 low expression. Based on this observation, we tried to classify luminal breast cancer samples

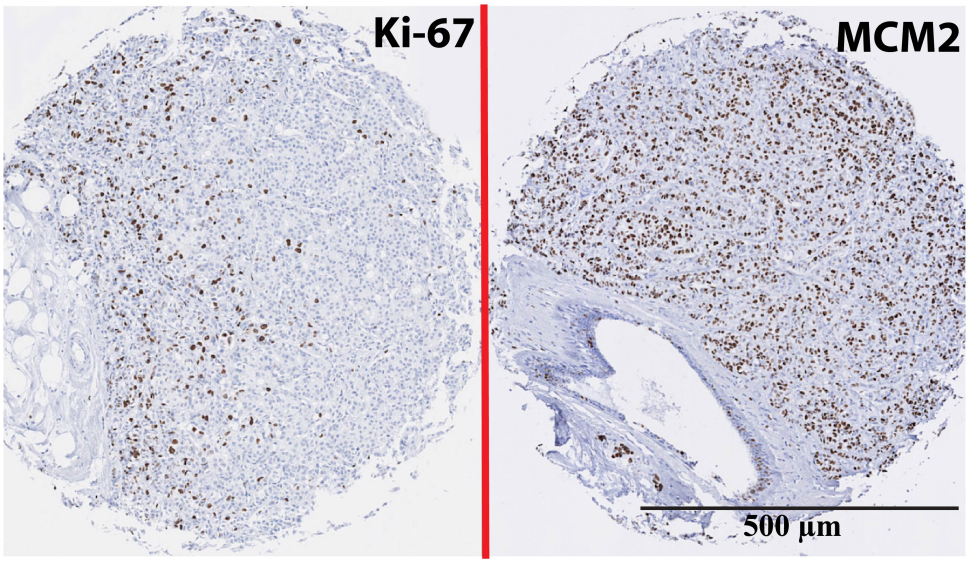


Figure 4.4: Intratumoral heterogeneity for Ki-67 labeling but not for MCM2, in breast cancer tissues.

Adjacent sections of the same breast cancer core stained with Ki-67 (on the left) and MCM2 (on the right). One observes conspicuous intratumoral heterogeneity for Ki-67 (only half of cells display nuclear labeling). However, no such heterogeneity was found with MCM2 on adjacent section of the same core (most of the cells are labeled with MCM2. Magnification 5X

Table 4.2: MCM2 visual and automated scores in different molecular subtypes of breast cancer

MCM2 Expression	Luminal A				Luminal B				HER2-positive				Triple-negative			
	Visual		Automated		Visual		Automated		Visual		Automated		Visual		Automated	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Low	103	95	100	93	10	21	12	25	12	19	17	27	27	19	27	19
High	5	5	8	7	38	79	36	75	52	81	47	73	117	81	117	81
Total	108	100	108	100	48	100	48	100	64	100	64	100	144	100	144	100

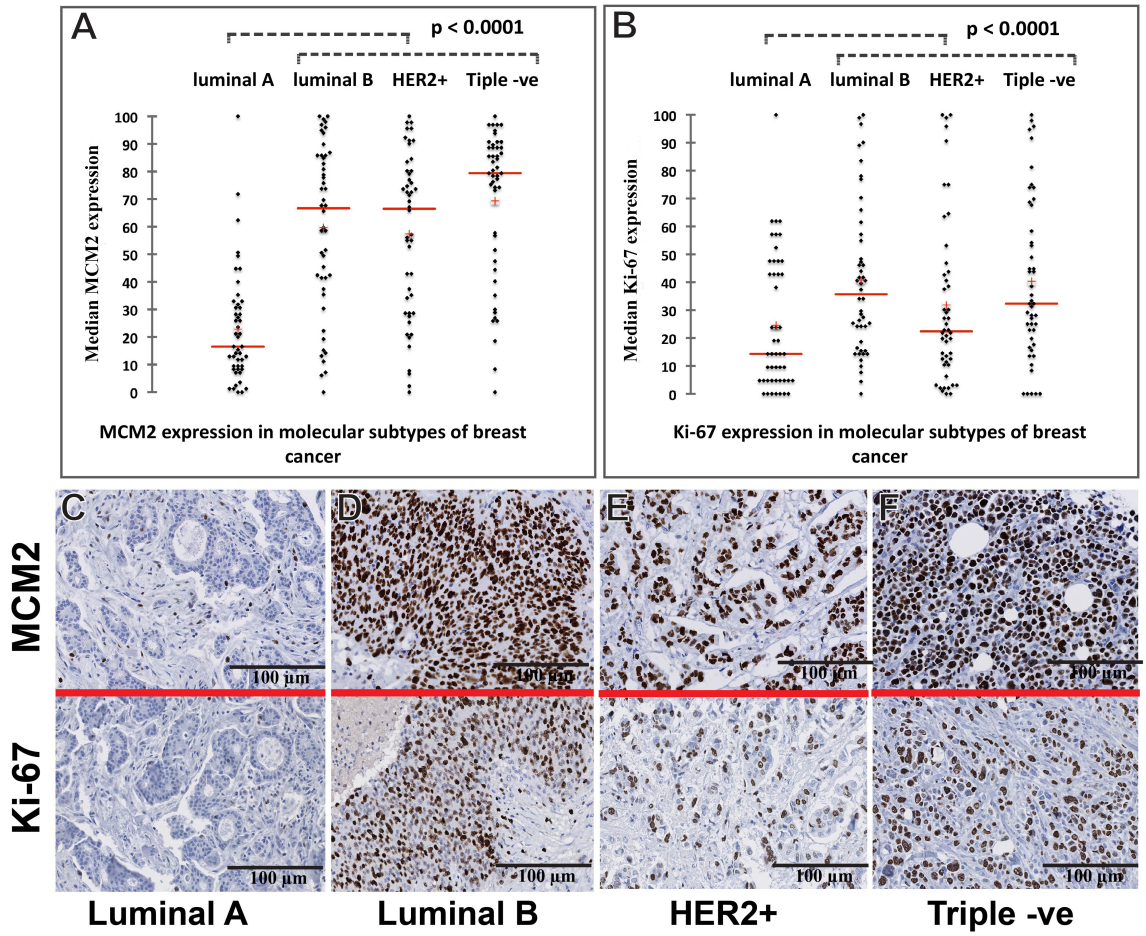


Figure 4.5: Overexpression of MCM2 is associated with triple-negative, luminal B and HER2-positive breast cancer

A) Elevated levels of MCM2 expression were detected in luminal B, HER2-positive and triple-negative breast cancer. Significant differences ($p < 0.0001$) were detected between these three subtypes and luminal A breast cancer. **B)** Significant differences ($p < 0.0001$) could also be detected between Ki-67 expression in luminal B, HER2-positive and triple negative breast cancer when compared to luminal A tumors. **C)** Luminal A, **D)** Luminal B, **E)** HER2-positive, **F)** Triple-negative breast cancer stained with MCM2 (upper panel), Ki-67 (lower panel). Overexpression of MCM2 and Ki-67 is associated with luminal B, HER2-positive and triple-negative breast cancer. MCM2 expression is consistently higher than of Ki-67 expression in different molecular subtypes of breast cancer. (Pictures were taken from adjacent sections of the same punches). Magnification 20X (C-F)

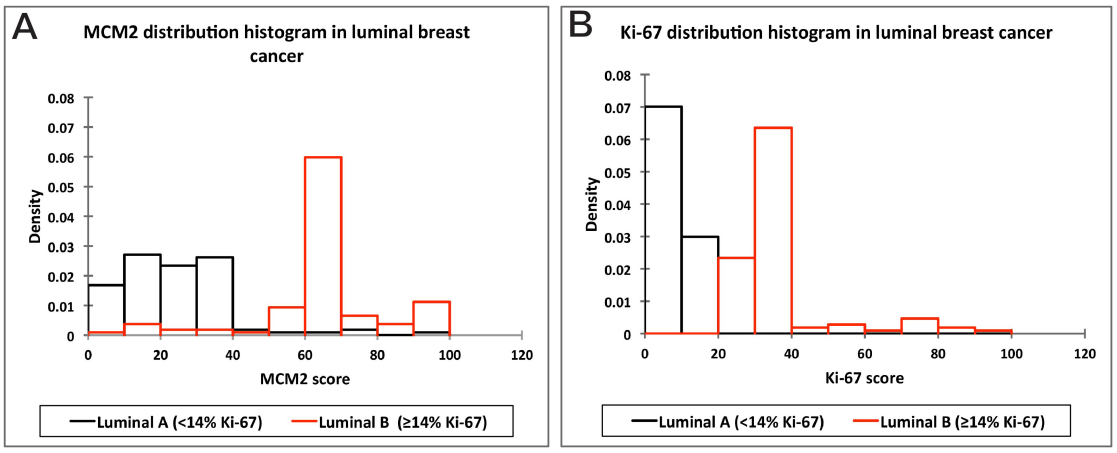


Figure 4.6: A bimodal frequency distribution of MCM2 scores is detected in luminal breast cancer

A) Frequency distribution analysis showing a bimodal distribution of MCM2 scores among luminal breast cancer. MCM2 labeling helps to distinguish two distinct subgroups in luminal breast cancers. Tumors with scores ranging from 0 to 38.4% are considered to represent tumors with low proliferative potential, while tumors whose scores are found to be between > 38.5-100% are considered to be tumors with high proliferative potential. **B)** When the same type of frequency distribution analysis was applied to Ki-67 scores, no such bimodal distribution pattern could be observed.

into luminal A and B based on the following criteria: luminal A (ER+, PR+, HER2-, MCM2 ≤ 40%), luminal B (ER+, PR+, HER2-, MCM2 > 40%). Our results indicate that if we were to apply a 40% index of MCM2 instead of 14% index of Ki-67 roughly 10% of tumors (15 out of 156) would have to be reclassified.

High levels of MCM2 are associated with a shorter latency to relapse

We next reviewed the clinical charts of 200 patients for the period extending from 2000 to 2013, to sort out the association between MCM2 and Ki-67 expression with clinical relapse. Our results demonstrate that enhanced expression of both of MCM2 and Ki-67 are associated with a shorter latency to clinical relapse in a statistically significant manner. However, the association between MCM2 expression and relapse (p = 0.0003) is tighter than that of Ki-67 (p = 0.0458) (Figure 4.7 A&B). We next conducted univariate Cox regression analysis on our data to determine the role of a number of factors such as level of MCM2, histological grades, histological subtypes and molecular subtype on the incidence of relapse for our patients. Results of the univariate analysis confirmed the association of MCM2 overexpression with a higher risk of relapse (HR = 6.98, 95%CI = 2.00-24.40, *p-value* = 0.002). Moreover, it shows that luminal A breast cancer has a significantly lower risk to develop relapse than HER2-positive breast cancer (HR = 0.11, 95%CI = 0.02-0.56, *p-value* = 0.01) (Table 4.3). However, no other factor was found to be significant in multivariate analysis.

Likewise, the Kaplan-Meier overall survival (OS) curve obtained from the same cohort of patients showed no significant association between both MCM2 and Ki-67 levels of expression and OS (Figure 4.7 C&D). The log-rank test demonstrated that the survival intervals are not significantly different in both groups at 5% level of significance, for MCM2 (p = 0.0631) and for Ki-67 (p = 0.453). It is interesting to note that the difference in survival after 5.76 years is almost 10% lower in patients with high MCM2 expression. The difference remains the same even after 10 years, indicating that patients who survive after 5.76 years maintain a constant risk of dying. Finally, no significant difference could be identified between survival after relapse (SAR) for patients having high and low levels of MCM2 (p = 0.0817) and Ki-67 expression (p = 0.355). Univariate and multivariate analyses confirmed the association between triple-

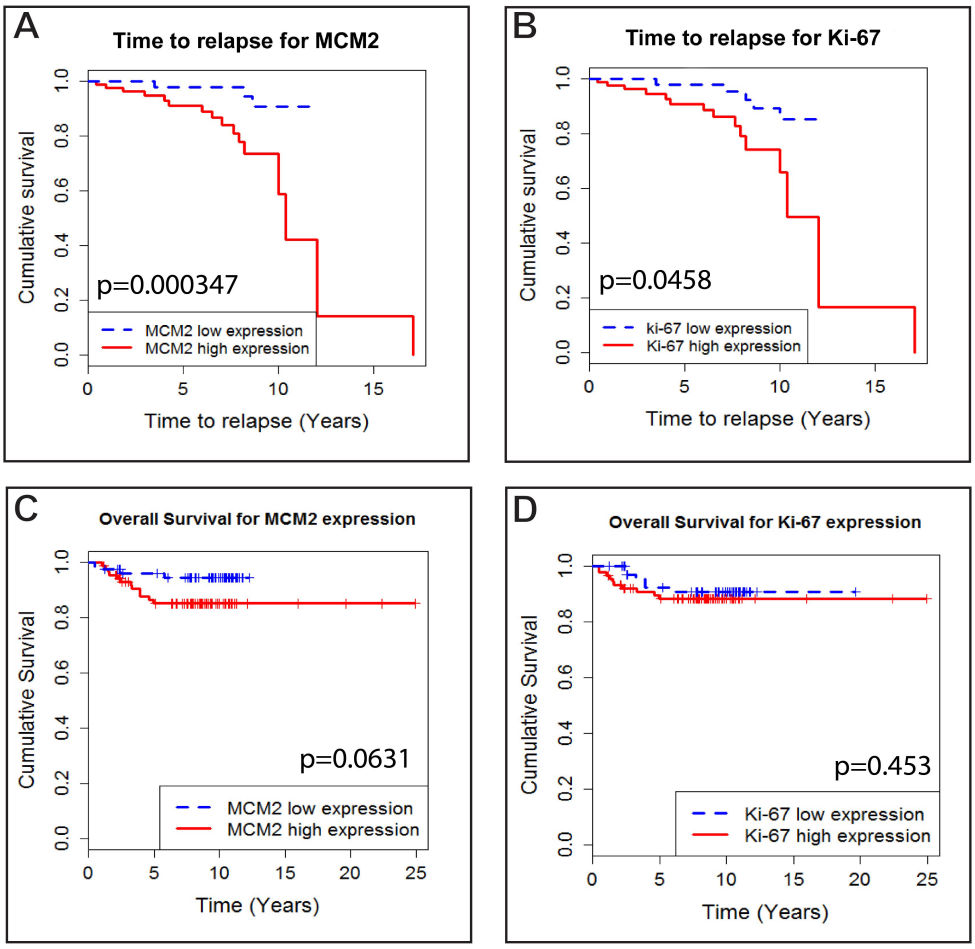


Figure 4.7: High levels of MCM2 are associated with a shorter latency to relapse

A) & **B)** Elevated levels of MCM2 and Ki-67 expression are associated with a shorter latency to clinical relapse. The association between MCM2 and relapse (p=0.0003) is tighter than that of Ki-67 (p= 0.0458). **C)** & **D)** No significant association between both MCM2 (p=0.0631) and Ki-67 (p=0.453) levels of expression and overall survival could be identified.

Table 4.3: Univariate Cox regression analysis of different factors that could affect OS, SAR and relapse.

	Relapse			**Overall survival (OS)**			**Survival after relapse (SAR)**		
Parameters	HR	95% CI	*p-value*	HR	95% CI	*p-value*	HR	95% CI	*p-value*
MCM2: High	**6.98**	**2.00-24.40**	**0.002**	2.43	0.77-7.64	0.128	2.43	0.75-7.35	0.14
Molecular subtypes									
HER2-positive	Reference			Reference			Reference		
Luminal A*	**0.11**	**0.02-0.56**	**0.01**						
Luminal B	0.94	0.27-3.22	0.92	0.87	0.25-2.97	0.82	0.97	0.28-3.31	0.96
Triple-negative	0.89	0.30-2.60	1.12	**0.23**	**0.07-0.78**	**0.0185**	**0.25**	**0.07-0.85**	**0.03**

HR= hazard ratio, CI= confidence interval. * No event occurred for Luminal A with OS and SAR

Table 4.4: **Multivariate Cox-regression analysis of different factors that could affect OS and SAR.**

	Overall survival (OS)			**Survival after relapse (SAR)**		
Parameters	HR	95% CI	*p-value*	HR	95% CI	*p-value*
MCM2: High	0.90	0.28-2.86	0.86	0.89	0.28-2.80	0.84
Molecular subtypes						
HER2-positive	Reference			Reference		
Luminal A*						
Luminal B	0.86	0.25-2.96	0.86	0.96	0.27-3.29	0.95
Triple-negative	**0.23**	**0.07-0.78**	**0.019**	**0.25**	**0.07-0.85**	**0.027**

HR= hazard ratio, CI= confidence interval, * No event occurred for Luminal A

negative breast cancer and risk of death (overall survival) (HR = 0.23, 95% CI=0.07-0.78, p-value=0.0185) and survival after relapse (HR = 0.25, 95%CI=0.07-0.85, p-value=0.03) in patients with high MCM2 expression. No other factors such as histological grades, histological subtypes and other molecular subtypes appeared to be significant in case of survivals (Tables 4.3&4).

Discussion

In the present study, we have compared the levels of expression of both MCM2 and Ki-67 as a means to assess cellular proliferation in a large cohort of patients with breast cancer of different grades and different molecular subtypes. We also compared the prognostic value of these two markers in predicting the overall survival, incidence of relapse and survival after relapse. Our findings confirmed that both MCM2 and Ki-67 are highly expressed in higher histological grades tumors especially in clinically aggressive breast cancers such as triple-negative, HER2-positive and luminal B subtypes tumors. Importantly, in contrast to Ki-67, two distinct subgroups among luminal breast cancer (ER+/HER2-) could be easily detected using a MCM2 labeling index of 40%.

Using TMAs comprising normal breast tissue and human breast cancers of different grades and subtypes, we found that MCM2 is present at higher levels than Ki-67 both in normal breast tissue and in different types of breast cancer (mean difference between the two scores reaching 40% in normal breast tissue and 25% in breast cancer $p < 0.0001$). This is consistent with previous observations by others reporting higher expression of MCM2 when compared to Ki-67 in normal breast tissues and breast carcinoma [344, 494]. In our opinion this indicates that MCM2 labeling is able to detect subsets of proliferating mammary epithelial cells that cannot be detected by Ki-67 alone. This is in line with the findings of Lopez et al. [305] who confirmed that Ki-67 is completely absent in the initial G1 phase of the cell cycle. Alternatively, one might surmise that Ki-67 protein is present inside the cells but cannot be detected due to pre-analytical condition such as fixation or because of altered biological properties of Ki-67 such as conformational changes or stable interactions with other proteins [347].

The dual pattern of MCM2 expression in normal breast tissue is intriguing. Whereas 62% of normal breast tissues (n =21) show MCM2 nuclear labeling in only a few scattered cells lining the terminal duct units, a significant proportion (36%) of normal tissues (n =21) display MCM2 expression in the vast majority of normal breast cells. The biological significance of these two sub-populations with distinct patterns of MCM2 expression is presently unknown and need further clarification. At this point, one can only hypothesize that the highly proliferative group reflects the state of hormonal stimulation in a given patient at the time of surgery. Whether or not it results in a higher susceptibility to neoplastic transformation is an unresolved question [495]. Clearly, more investigations are needed with larger cohorts of normal breast tissues and their follow up data to clarify this issue.

Our results also indicate that both MCM2 and Ki-67 are highly expressed in breast cancer of higher histological grades. This is in agreement with Ali et al. who reported a significant correlation between proliferative markers including both MCM2 and Ki-67 and breast cancer grades [496]. Our findings coming from both *in silico* analyses and immunohistochemistry also support the notion that Ki-67 and MCM2 are highly expressed both at the mRNA and protein levels in subgroups of triple-negative and HER2-positive as well as in luminal B subtypes of breast cancer [497-500]. Indeed, our data revealed that MCM2 expressed at higher levels than Ki-67 in aggressive molecular subtypes of breast cancer such as triple-negative, HER2-positive and luminal B breast cancer when compared to luminal A breast cancer. To our knowledge, there has been no previous report in the literature that specifically correlated MCM2 expression with individual breast cancer molecular subtypes.

Based on Ki-67 labeling distribution, it has been previously proposed that ER-positive breast cancer form a continuum rather than segregating into distinct subtypes [23]. Our data presented herein clearly challenge this interpretation showing that two distinct subgroups of hormonally responsive breast cancers (ER+/HER2-) can be identified based on the estimated percentage of MCM2 positive tumor cells using a threshold of 40% set in accordance with the ROC curve. Hence, ER+/HER2- tumors with MCM2 expression ranging from 0-40% are considered to be luminal breast cancer with low proliferative potential, while ER+/HER2- tumors with over expression of MCM2

(>40%) are considered to represent luminal breast cancer with high proliferative potential. The fact that we observed a convincing bimodal distribution of MCM2 in luminal breast cancer along with a recent gene expression study reporting that MCM2 but not Ki-67 belongs to a class of breast cancer genes with bimodal distribution [501] supports our views.

Lastly, we sought to correlate the expression profiles of MCM2 and Ki-67 with clinical outcome characteristics such as survival rates and tumor relapse. Our results confirmed that overexpression of both markers is associated with shorter latency to clinical relapse. However, the association between MCM2 and relapse ($p = 0.000347$) is more tightly correlated than that of Ki-67 ($p = 0.0458$). Our data strongly support the role of Ki-67 and MCM2 as prognostic markers in breast cancer [50, 344]. One conceivable explanation to the poor outcome associated with high MCM2 expression is the flawed DNA replication licensing system, MCM2-MCM7, which leads to uncontrolled cell proliferation [502].

There are many reasons to believe that MCM2 outperforms Ki-67 as a tool to assess cell proliferation in breast cancer. Firstly, in contrast to Ki-67, MCM2 can label all proliferative cells during the active phases of the cell cycle, disappearing only when cells are quiescent [503]. Moreover, Ki-67 labeled only a fraction of proliferating cells entering the G1 phase of the cell cycle for the onset of sustained Ki-67 expression occurs only in late G1 [305, 306]. Secondly, Ki-67 labeling (clone SP6, BioCare medical) suffers from both intra-and inter-sample heterogeneity as well as unpredictable changes in labeling intensity, a phenomenon that has not been observed to the same extent with MCM2 staining. One plausible explanation to the variability of Ki-67 staining may be related to the normal variation of Ki-67 concentration during different phases of the cell cycle with accumulation during G2 and S phases and a nadir during anaphase and telophase of the mitotic cycle [307-309]. One should also bear in mind that the half life of Ki-67 is estimated to be approximately 90 minutes followed by its rapid degradation [504]. Evidently, one does not expect such variability in staining intensity with MCM2 as it remains confined to the nuclei at all times exerting its function through reversible binding to the chromatin instead of undergoing alternating cycles of synthesis and degradation as Ki-67 [334]. Thirdly, the role of MCM2 in DNA replication is now firmly

established, as eukaryotic cells cannot proliferate in the absence of MCM2 [334]. Although Ki-67 plays an important role in cell division its mechanism of action inside the cell is poorly understood [281]. Notably cells that are depleted of Ki-67, maintain a normal cell-cycle profile albeit their smaller nuclei [310].

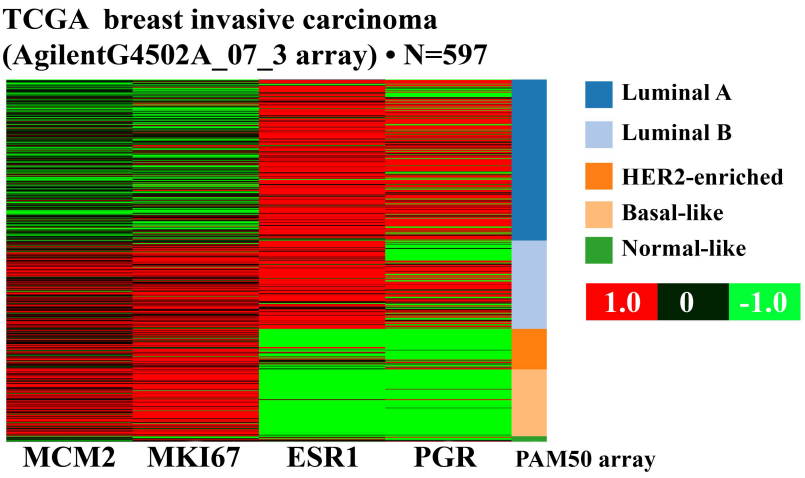
Even though we feel that the data presented in this paper are quite compelling, it is fair to say that there are a few intrinsic limitations to our study. Firstly, our definition of luminal A and luminal B tumors was based on the following surrogate markers; ER, PR, HER2 and 14% index of Ki-67. Assessment of MCM2 expression in these two groups defined by mRNA expression should clearly be performed as a follow-up study. Secondly, our results, including the scoring methods (visual and automated) and the proposed cutoff point, need to be validated on another independent cohort of patients. Thirdly, follow up data on a new set of patients is required to independently confirm our findings. Lastly, whole sections rather than TMA cores are needed to check on the robustness of MCM2 especially in regions found to be heterogeneous using Ki-67 staining.

Conclusion

In conclusion, our results indicate that MCM2 outperforms Ki-67 as a proliferative and prognostic marker of breast cancer. Although Ki-67 revealed itself an independent prognostic marker, many practical issues are associated with its clinical use: tumor heterogeneity often observed in cancers, the unyielding 14% threshold, and the fact that some proliferating cells escape Ki-67 labeling. In contrast, MCM2 seems to overcome many of these shortcomings as it detects all proliferating cells including those that are not detected by Ki-67. It also provides similar or even better prognostic information more likely to be clinically applicable because of a more convenient to use 40% cutoff point. It is our opinion that, MCM2 has the ability to classify ER+/HER2- breast cancer into two subsets of hormone sensitive tumors with diverging proliferative status. It is hoped that such defined subsets of tumors will facilitate the management of hormone sensitive breast cancers.

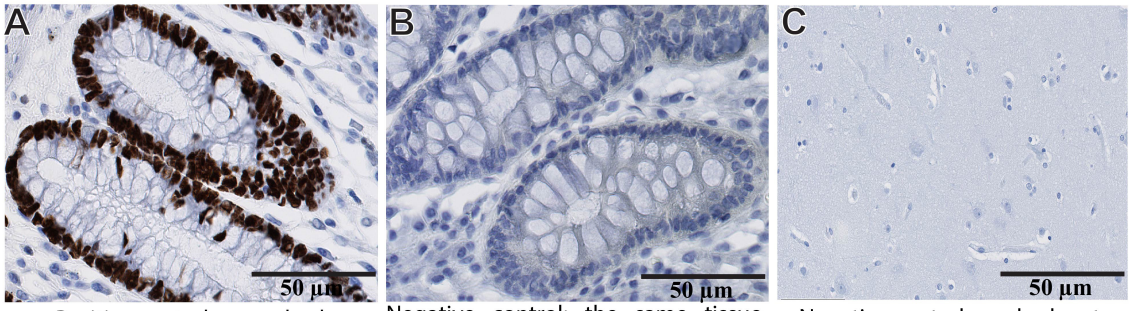
Additional materials

Additional figures:



Additional file 1: *In silico* analysis of *MCM2*, *MKI67, ESR1* and *PGR* mRNA expression in PAM50 molecular subtypes of 597 breast cancers obtained from UCSC Cancer Genomics Browser

The heat map exhibits the expression of *MCM2, MKI67, ESR1 and PGR* in different molecular subtypes of breast cancer. Red represents data values > 0, green represents values < 0 and black represents value = 0. *MCM2 and MKI67* are highly expressed in basal-like, HER2-enriched and luminal B breast cancer while low levels of expression were detected in luminal A and normal-like subtypes.



Positive control: normal colon

Negative control: the same tissue incubated with non-immune serum containing IgG (same isotype/same species)

Negative control: cerebral cortex

Additional file 2: Validation of MCM2 antibody specificity for IHC studies.

A) Normal human colonic mucosa showing an intense nuclear labeling of the epithelial cells lining the colonic glands after incubating the section with an anti-MCM2 primary antibody (positive control). **B)** Adjacent section from the same colonic tissue incubated with a non-immune serum (IgG same isotype/ same species) showing a complete lack of MCM2 labeling (negative control). **C)** Human cerebral cortex incubated with an anti-MCM2 primary antibody showing complete lack of MCM2 expression (negative control). Magnification 40X

Additional tables:

Additional file 3: Immunohistochemical scoring systems

Marker	Subcellular localization	Scoring system	Total Score	Criteria	Results
ER and PR	Nuclei	Allred score	0-8	Sum of the proportion and average intensity scores of positive tumor cells	Negative = 0-2 Positive = 3-8
HER-2	Membrane	CAP-approved scoring system	0, 1+, 2+, 3+	0= No immunostaining or membrane staining which is incomplete or barely perceptible within ≤ 10% of the invasive tumor cells. 1+= Incomplete membrane or barely perceptible staining within >10% of invasive tumor cells 2+= Circumferential membrane staining that is incomplete and/or weak/moderate within >10% of the invasive tumor cells or complete membranous staining that is intense within ≤ 10% of the invasive tumor cells 3+= Circumferential membranous staining that is complete and intense in >10% of tumor cells	0-1+ are negative 2+ is equivocal 3+ is positive
Ki-67	Nuclei	% of positively	0-100%	% of positively stained nuclei	Ki-67 low = 0-<14%

		stained nuclei			Ki-67 high= ≥14-100%
MCM2	Nuclei	% of positively stained nuclei	0-100%	% of positively stained nuclei	MCM2 Low = 0-40% MCM2 High >40-100%

Additional file 4: Classification of molecular subtypes of breast cancer based on IHC surrogate markers [112, 421]**.**

Molecular subtypes	IHC expression of surrogate markers
Luminal A	ER positive and/or PR positive HER2 negative Ki-67 <14%
Luminal B	A. ER positive and/or PR positive HER2 negative Ki-67 ≥14% B. ER positive and/or PR positive HER2 positive Any Ki-67
HER2-positive	ER negative PR negative HER2 positive Any Ki-67
Triple negative	ER negative PR negative HER2 negative Any Ki-67

CHAPTER V

Discussion

During the past decade, pioneering unsupervised gene expression profiling studies on invasive breast cancers have identified different molecular subgroups with distinct patterns of gene expression: luminal A, luminal B, normal breast-like, HER2-positive and basal-like breast cancer [36]. Breast tumors of the luminal subtypes express genes characteristic of luminal epithelial cells whereas the basal-like breast cancers express genes that are normally found in normal basal and/or myoepithelial cells [36]. Amongst the two hormonally responsive tumors types: luminal A and luminal B tumors there are significant biological and clinical differences. Indeed, it is known that luminal B breast cancer are associated with an overall worse prognosis than luminal A tumors, akin with two other notoriously aggressive breast cancers: HER2-enriched and basal-like subtypes [41, 49]. Because of the lack of actionable therapeutic targets, basal-like breast carcinomas have been associated with a worse prognosis, higher rates of recurrence and shorter survival when compared to other types of breast cancer [64, 66]. The high mortality rate of basal-like breast cancer is reflected by the higher incidence of visceral and brain metastases [65, 505]. Lack of expression of the three main surrogate breast markers ER, PR and HER2 has been used to operationally define basal-like breast cancer also known as triple-negative breast cancer [22].

During the past several years, it has become apparent that invasive breast cancers of similar clinical stage and histological grade may behave differently not only in terms of clinical outcome but also in their ability to respond to treatment [156]. This emphasizes the need for new prognostic biomarkers of breast cancer that can better stratify patients with breast cancers. The fundamental aim of the prognostic biomarkers is to give an idea about the behavior of the tumor, regardless of the type of treatment, such as its aggressiveness, probability of its recurrence and the risk of developing metastases with the highest possible sensitivity and specificity [126, 135]. Our main hypothesis is that differential patterns of gene expression in breast cancer can help distinguish

subgroups of tumors with specific clinical outcomes and response to treatment. During the course of my doctoral studies, my aim has been to identify and characterize new biomarkers that could best define aggressive subtypes of breast cancer focusing primarily on triple-negative cancers. A second albeit related aim was to identify novel biomarkers that could help distinguish luminal A from luminal B breast cancer.

To achieve those aims, I first conducted *in silico* analysis studies on available DNA microarray and RNA sequencing data of human breast cancer tissues and breast cancer cell lines. The *in silico* analyses helped us identify two types of genes: those that are differentially expressed in triple-negative breast cancer and those that are significantly associated with luminal B breast cancers but not with luminal A subtype (unpublished results, see Appendix IV). Evidently, we had to limit ourselves to the markers that appeared more robust or, at least, promising such as ANXA1, MMP-9 and DP103 for triple-negative breast cancer and MCM2 for luminal breast cancer. To that end, *in silico* analysis proved invaluable to study the expression of genes of interest found in different molecular subtypes of breast cancer and to correlate their expression with that of clinically relevant markers such as ER, PR and HER2. With this information at hand we constructed seven tissue microarrays (TMAs) from paraffin blocks containing human breast cancers and normal breast tissues obtained at surgery. We also thought fit to include cell microarrays (CMAs) from a panel of human breast cancer cell lines with varying degree of ER, PR and HER2 expression. In all cases, levels of protein expression were detected using automated immunohistochemistry and compared to that of other conventional breast cancer markers (ER, PR, HER2, Ki-67, CK 5/6, VIM, FOXA1). Lastly, our IHC assays were correlated with patients' clinical data (metastasis, relapse and survival rates). All our aims and objectives were accomplished, with the scientific relevance and importance of our findings being discussed in each related chapter (II, III & IV). However, in this discussion, I will focus on the main results of our biomarkers and discuss their clinical significance as potential biomarker in breast cancer.

I. Novel biomarkers in triple-negative molecular subtype of breast cancer

Differential expression of ANXA1 has been detected in normal human breast tissue and in breast cancers. ANXA1 expression is severely deregulated in high-grade

breast cancers with poor clinical outcomes. About half of the samples collected from patients with triple-negative and 10% of HER2-positive breast cancers had strikingly elevated levels of ANXA1 expression. Moreover, human triple-negative breast cancer cell lines expressed ANXA1 at significantly higher levels when compared to all other groups. This supports similar findings by Yom et al. who demonstrated that ANXA1 expression was significantly correlated with hormone receptor negativity, HER2-positivity and triple-negative breast cancers [379]. In addition, the association of ANXA1 with basal-like breast cancer, which is considered to be a subset of triple-negative breast cancer, was further established using concomitant basal cytokeratin CK5/6. All in all, these results are in agreement with those of Perou et al. who demonstrated an association between the ANXA1 gene and basal subtype phenotype of breast cancer [36].

Notably, we also demonstrated that ANXA1 expression is normally confined to the myoepithelial cell layer surrounding ducts and lobules present in normal breast tissue and *in situ* carcinomas. Again, our results are consistent with those of Ang et al. who reported a similar pattern of expression of ANXA1 in myoepithelial cells [359]. Although we were tempted to raise the possibility that basal-like breast cancer originate from myoepithelial cells, findings by Livasy et al. failed to support a myoepithelial origin for basal-like breast cancers as they completely fail to express other myoepithelial markers such as P63 and αSMA [100]. Furthermore, constitutive expression of ANXA1 in myoepithelial cells provided a likely explanation for the high level of ANXA1 expression found by *in silico* analysis in normal-like breast cancer subtype given the high degree of contaminating normal breast tissue in these tumor samples [426]. Of note, our results indicated a complete lack of ANXA1 expression in all luminal A and luminal B breast cancer tissue as well as luminal cell lines. Taken together, these results indicate the ability of ANXA1 to dissect out subsets of high histological grades breast cancers and help provide a better understanding of breast cancer heterogeneity.

Several studies have emphasized a potential role of ANXA1 in enhancing the EMT and increasing metastatic potential of breast cancer cells [374, 425]. For example, de Graauw indicated that ANXA1 promotes metastasis development in basal-like breast cancer cells by enhancing TGFβ/SMAD signaling and actin reorganization [375]. Based

on *in silico* analysis and IHC, our results, suggested a positive correlation between ANXA1 and EMT associated markers including VIM, MSN and LYN. We therefore suspected a potential role for ANXA1 in enhancing EMT and breast cancer cell metastasis. However, to our surprise, others have shown that ANXA1 attenuates EMT and inhibits metastasis of breast cancer [108, 376]. Further experiments are warranted to resolve this issue and explore the role of ANXA1 in breast cancer progression and further development of metastasis.

Previous studies have established a role for MMP-9 in several key processes that contribute to breast cancer development and progression including cancer cell migration, invasion and metastasis [393, 395-400]. In line with recent studies by others, our results have demonstrated a significant increase in MMP-9 expression in breast cancer cells where compared to normal breast tissue. Notably, we have established that elevated levels of MMP-9 are closely associated with breast cancers of high histological grade, which has been supported by recent literature [506]. High level of MMP-9 was detected in 79.4% of triple-negative and 87.9% of HER2-positive breast cancer tissues and in all samples of lymph nodal metastases. Again, our results are in agreement with that of Zhao et al. who demonstrated that high MMP-9 expression is correlated with triple negative breast cancers [403]. Of note, La Rocca reported elevated serum level of MMP-9 in HER2-positive subtypes of breast cancers using gelatin zymography providing additional evidence on the role of MMP-9 [404].

It is well documented that both triple-negative and HER2-positive breast cancer molecular subtypes are associated with aggressive behavior and high metastatic potential [55, 507]. Our results provide strong evidence that increased expression of MMP-9 contribute to breast cancer cell dissemination and poor clinical outcome. In our opinion, MMP-9 overexpression could help segregate subsets of aggressive breast cancer into more clinically meaningful subtypes. This is substantiated by our data confirming that overexpression of MMP-9 is tightly correlated with lymphovascular invasion, regional node metastasis, a shorter time to relapse and a reduced SAR. Interestingly, our findings are also supported by independent work by others who also established that up-regulation of MMP-9 is associated with metastasis, lymphatic invasiveness, poor prognosis and unfavorable clinical course of breast cancer [402, 403, 464]. Taken together, several

lines of investigations underpinned the contribution of MMP-9 in promoting breast cancer metastasis and collectively highlighted the clinical relevance of MMP-9 overexpression in subsets of breast cancer. In addition, MMP-9 has been included among the MammaPrint® 70 genes panel, the first fully commercialized and FDA approved microarray-based multigene assay for breast cancer [272, 465]. Hence, we strongly believed that MMP-9 overexpression is sufficiently compelling to warrant clinical studies to validate its clinical utility as a breast cancer prognostic biomarker. Furthermore, one might consider including MMP-9 alone or in combination with other genes in the development of other multigene multiplex assays.

II. Novel biomarkers to distinguish luminal A and luminal B molecular subtypes of breast cancer

Estrogen receptor positive (ER+/HER2-) breast cancers comprise two distinct subtypes; luminal A and luminal B. Each of them differs from one another by clinical outcomes, biological characteristics such as proliferation and response to treatment [41, 44, 45]. Because Ki-67 assessment has raised a number of practical issues, it has failed to impose itself as a routinely utilized proliferative marker to discriminate between luminal A and luminal B molecular subtypes of breast cancer [45, 50, 281, 282, 327-330]. Since cell proliferation is known to be the single most important parameter apart from ER, PR and HER2 in luminal breast cancers, there is a critical need for a robust and validated immunohistochemical assay to distinguish between luminal A and luminal B tumors. Interestingly, MCM2, whose role in DNA replication and cell proliferation is now firmly established, has been reported to be a promising proliferative marker in many different types of cancer [283, 483-485].

Our results indicate that MCM2 is highly expressed in breast cancer of high histological grades that comprise clinically aggressive tumors such as triple-negative (81%), HER2-positive (81%) and luminal B (97%) subtypes of breast cancer. Our results confirm findings by Ali et al. who reported a significant correlation between MCM2 and Ki-67 in high-grade breast cancers [496]. However, to our knowledge, there has been no previous report in the literature that specifically correlated MCM2 expression with individual breast cancer molecular subtypes. In agreement with the work of others [344,

494], our findings demonstrate that MCM2 is often expressed at significantly higher levels than Ki-67 in normal breast tissues and in breast cancers. We believe that MCM2 labeling is able to detect subsets of proliferating mammary epithelial cells that cannot be detected by Ki-67 alone. Alternatively, one might assume that Ki-67 protein is present inside the cells but cannot be detected due to pre-analytical condition such as fixation or because of altered biological properties of Ki-67 such as conformational changes or stable interactions with other proteins [347].

During the course of our investigation on luminal (ER+/HER2-) breast cancers we found out a bimodal frequency distribution of MCM2 scores with regards to luminal breast cancers. This distribution pattern agrees with that of Bessarabova et al. who demonstrated that MCM2, but not Ki-67, was one of the bi-modally expressed genes based on published dataset of multiple microarray platforms [501]. Of note, two distinct subgroups among hormonally responsive luminal breast cancer could be easily identified using a 40 % index of MCM2.

In our hands, MCM2 provided similar or even better prognostic information than Ki-67. On the one hand, an enhanced expression of both of MCM2 and Ki-67 were found to be associated with a shorter latency to clinical relapse in a statistically significant manner. Moreover, MCM2 expression and relapse is more tightly associated than with of Ki-67. On the other hand, high expression of MCM2 is associated with shorter overall survival than that of Ki-67 but neither of them reached statistical significance. The above-mentioned findings support the results of two clinical investigations that stressed the potential role of Ki-67 and MCM2 as prognostic markers in breast cancer, which highlight the importance of extending the analysis of MCM2 on new cohorts of patients [50, 344].

The dual pattern of MCM2 expression in normal breast tissue is intriguing. Whereas 62% of normal breast tissues (n =21) show MCM2 nuclear labeling in only a few scattered cells lining the terminal duct units, a significant proportion (36%) of normal tissues (n =21) display MCM2 expression in the vast majority of normal breast cells. The biological significance of these two sub-populations with distinct patterns of MCM2 expression is presently unknown and need further clarification. At this point, one can only hypothesize that the highly proliferative group reflects the state of hormonal

stimulation in a given patient at the time of surgery. Whether or not it results in a higher susceptibility to neoplastic transformation is an unresolved question [495]. Clearly, more investigations are needed with larger cohorts of normal breast tissues and their follow up data to clarify this issue.

In the present work, we provided strong evidence that Ki-67 can be advantageously substituted by MCM2 to measure cell proliferation in breast cancer, as it seems to overcome many of Ki-67 shortcomings. Firstly, MCM2 can label all proliferative cells since it is present at all stages of the cell cycle and disappears only when cells become quiescent. Secondly, variability in staining intensity was minimal and, so far, no heterogeneity could be detected using MCM2 labeling. Lastly and more importantly, a bimodal distribution of MCM2 labeled luminal tumors could be observed using a 40% MCM2 threshold suggesting that one can readily distinguish two distinct groups of patients whose treatment and clinical outcomes are likely to diverge.

The use of *in silico* analyses in our work was a powerful approach for identifying candidate biomarkers with potential clinical significance in breast cancer. In addition, combining IHC together with TMA was a very efficient high throughput methodology that helped us assess different biomarkers on a large number of tissue samples simultaneously. Using IHC can be considered as one of the most readily applicable techniques for routine clinical practice as it is already used for ER, PR and HER2. Furthermore, it has been validated in different studies that IHC define clinically and biologically relevant breast cancer subgroups comparable to that identified by gene expression profiling [110].

However, many limitations have been reported with using IHC such as the technical reproducibility and the interpretation and quantifications of the obtained expression [108, 109]. During our work, we tried to overcome these issues by standardization of various aspects of IHC techniques such as using comparable experimental conditions during slide preparations, automated IHC staining and whole slide scanning. Also, to reduce the subjectivity in interpreting and quantifying biomarker expression, we used two different scoring methods: a visual method using conventional light microscopy and a computer-assisted automated scoring method using Visiomorph© and Tissuemorph© Digital Pathology (DP) softwares. Although visual scoring of IHC

reactions is labor-intensive, time consuming and is subject to intra-observer and inter-observer variability, two independent observers performed the evaluations to increase the reproducibility of this scoring system. Using the automated scorings resulted in a much more rapid and accurate readout. Of note, assessment of concordance between the two scoring systems to compare visual and automated scores and to evaluate their relative performance was carried out. Finally, determining the cutoff points of high and low expression was very tricky with the first biomarkers (ANXA1 and MMP-9). However, with MCM2 we tried to overcome this limitation by applying the ROC curve on the continuous automated scores to precisely detect the optimal cutoff point.

That being said, there are still a number of shortcomings that we cannot ignore in our studies. Firstly, using TMA allowed us to study only small portions of tissues from carcinoma. Due to the heterogeneity of breast cancer, using only small cores of tissue may not accurately match the expression of certain biomarker or for that matter the proliferative activity of cancer cells. Although excellent concordance between TMAs and complete section has been reported [508], whole sections rather than TMA cores are required to check on the robustness of our biomarkers expression. Secondly, our definition of luminal A and luminal B tumors was based on the following surrogate markers: ER, PR, HER2 and 14% index of Ki-67. This is hard to reconcile with our work on MCM2 as we demonstrated that Ki-67 is not an ideal proliferative biomarker in breast cancer. Using DNA microarray analysis could be a good way to better determine different molecular subtypes of breast cancer. Thirdly, single antibody for each biomarker has been tested and we are not absolutely certain that our results would be similar using another panel of antibodies. Lastly, our results need to be validated on an independent cohort of patients to check on the robustness and reproducibility of our findings.

Conclusion

Taken together, the experimental work presented herein revealed that ANXA1, MMP-9 and MCM2 are new biomarkers that can help dissect out subsets of aggressive breast cancer into clinically meaningful subtypes. Not only they can facilitate the identification and further characterization of unique and specific regulatory pathways, they also lead to a better understanding of the pathophysiological mechanisms involved in tumor progression and metastasis. If the clinical validity and utility of these markers can be confirmed on independent cohorts of breast cancer patients, they will have a significant impact on clinical practice and precision medicine. In conclusion, ANXA1, MMP-9 and MCM2 are valuable gene/protein candidates to be used alone or in combination with other genes in the development of a multi-gene panel or multiplex proteomic assay to predict clinical outcome.

Perspectives

This work emphasizes the need for further investigation into different molecular subtypes of breast cancer especially the aggressive subtypes. In addition, it highlights the importance of developing new prognostic biomarkers in breast cancer as this provides a more comprehensive picture of breast cancer heterogeneity and helps identify new targeted therapies specific for the different groups of breast cancer. Additional questions have been proposed to continue this work in the future including the following:

A) Development of multiplex IHC-based proteomic assay:

In the future, accurate breast cancer diagnosis and patient stratification will require the analysis of multiple biomarkers rather than the use of single gene assays [509]. The combined use of several markers such as ANXA1, MMP-9, MCM2 and DP103 may be a better tool than to use each candidate molecule individually in order to best stratify triple-negative breast cancers. A new cohort of patients with complete clinical follow up data, should make it possible to determine the usefulness of assessing the combined expression of these biomarkers.

B) Future work for MCM2:

Two important issues have been proposed to continue our work on MCM2; the first one concerns the possibility of using MCM2 expression to detect breast cancer recurrence instead of using the Oncotype Dx ® (RT-PCR based, Recurrence Score Assay). The second question revolves around the role of MCM2 expression in predicting clinical response in the context of neoadjuvant chemotherapy. To answer these questions, we need to construct new TMAs from a cohort of patients preferably those with ER+, node negative tumors that already had an Oncotype Dx assay and who are candidate for neoadjuvant chemotherapy. Follow up data (relapse, survival and metastasis) on this new cohort of patients should be available for review. Furthermore, DNA microarrays will be used to identify the luminal subtypes of breast cancer based on expression profile of a cluster of cell proliferation associated genes.

MCM2 need to be further investigated in a new cohort of breast cancer patients by IHC to confirm the reproducibility of a 40% cutoff point. We also need to check on the MCM2 status on patients whose recurrence scores are available. Finally, we need to compare the expression of MCM2 in breast cancer tissues from the same patients before and after receiving neoadjuvant chemotherapy.

c) Future work for other proliferative markers:

By continuing the work on other proliferation associated markers already identified by *in silico* analysis (Appendix IV) it may be possible to develop a protein/s based assay, which precisely defines the proliferation status of breast cancer patients.

APPENDIX

Appendix I: Panel of genes in MammaPrint assay (70-gene assay) [272].

Groups of genes	Gene name	Gene description
Genes regulating cell cycle (15 genes)	CCNE2	Cyclin E2
	ECT2	Epithelial cell transforming sequence 2 oncogene
	CENPA	Centromere protein A
	LIN9	lin-9 homolog
	KNTC2	Kinetochore associated 2
	MCM6	Minichromosome maintenance complex component 6
	NUSAP1	Nucleolar and spindle associated protein 1
	ORC6L	Origin recognition complex, subunit 6 like
	TSPYL5	TSPY-like 5
	RUNDC1	RUN domain containing 1
	PRC1	Protein regulator of cytokinesis 1
	RFC4	Replication factor C 4, 37 kDa
	RECQL5	RecQ protein-like 5
	CDCA7	Cell division cycle associated 7
	DTL	Denticleless homolog
Genes regulating cell growth and apoptosis (9 genes)	ESM1	Endothelial cell-specific molecule 1
	IGFBP5	Insulin-like growth factor binding protein 5
	FGF18	Fibroblast growth factor 18
	SCUBE2	Signal peptide, CUB domain, EGF-like 2
	TGFB3	Transforming growth factor, beta 3
	WISP1	WNT1 inducible signaling pathway protein 1

	TGFB3	Transforming growth factor, beta 3
	BBC3	BCL2 binding component 3
	EGLN1	egl nine homolog 1
Genes regulating cell proliferation (12 genes)	FLT1	fms-related tyrosine kinase 1
	HRASLS	HRAS-like suppressor
	STK32B	Serine/threonine kinase 32B
	RASSF7	Ras association (RalGDS/AF-6) domain family 7
	DCK	Deoxycytidine kinase
	MELK	Maternal embryonic leucine zipper kinase
	EXT1	Exostoses 1
	GNAZ	Guanine nucleotide binding protein, alpha z polypeptide
	EBF4	Early B-cell factor 4
	MTDH	Metadherin
	PITRM1	Pitrilysin metallopeptidase 1
	QSCN6L1	Quiescin Q6-like 1
Genes controlling cell invasion and metastasis (8 genes)	COL4A2	Collagen, type IV, alpha 2
	GPR180	G protein-coupled receptor 180
	MMP9	Matrix metallopeptidase 9
	GPR126	G protein-coupled receptor 126
	RTN4RL1	Reticulon 4 receptor-like 1
	DIAPH3	Diaphanous homolog 3
	CDC42BPA	CDC42 binding protein kinase alpha
	PALM2	Paralemmin 2
Genes controlling adaptation to unfamiliar microenvironment (7 genes)	ALDH4A1	Aldehyde dehydrogenase 4 family, member A1
	AYTL2	Acyltransferase like 2
	OXCT1	3-oxoacid CoA transferase 1, nuclear gene encoding mitochondrial protein
	PECI	
	GMPS	

	GSTM3	Peroxisomal D3,D2-enoyl-CoA isomerase
	SLC2A3	Guanine monphosphate synthetase
		Glutathione S-transferase M3
		Solute carrier family 2, member 3
Genes controlling angiogenesis (6 genes= 3 new + 3genes mentioned before)	**FLT1**	fms-related tyrosine kinase 1
	FGF18	Fibroblast growth factor 18
	COL4A2	Collagen, type IV, alpha 2
	GPR180	G protein-coupled receptor 180
	EGLN1	egl nine homolog 1
	MMP9	Matrix metallopeptidase 9
Miscellaneous genes (16 genes)	LOC100288906	Hypothetical protein LOC100288906
	C9orf30	Chromosome 9 open reading frame 30
	ZNF533	Zinc finger protein 533
	C16orf61	Chromosome 16 open reading frame 61
	SERF1A	Small EDRK-rich factor 1A
	C20orf46	Chromosome 20 open reading frame 46
	LOC730018	Similar to hCG1980668
	LOC100131053	Hypothetical LOC100131053
	AA555029_RC	No significant similarity found
	LGP2	Likely ortholog of mouse D11lgp2
	NUM	Neuromedin U
	UCHL5	Ubiquitin carboxyl-terminal hydrolase L5
	JHDM1D	Jumonji C domain containing histone
	AP2B1	demethylase 1 homolog D
	MS4 A7	Adaptor-related protein complex 2, beta 1
	RAB6B	subunit
		Membrane-spanning 4-domains, subfamily A, member 7
		RAB6B, member RAS oncogene family

Appendix II: Panel of genes in PAM50 Breast Cancer Intrinsic Classifier™ assay (50-gene assay) [276].

Gene name	Gene description
CDC6	Cell Division Cycle 6
CDC20	Cell Division Cycle 20
PGR	Progesterone Receptor
EGFR	Epidermal growth factor receptor
ESR1	Estrogen receptor 1
MKI67	Marker of proliferation Ki-67
UBE2T	Ubiquitin-conjugating enzyme E2T
BIRC5	Baculoviral IAP repeat containing 5
NUF2	NUF2, NDC80 Kinetochore Complex Component
TYMS	Thymidylate synthetase
MYBL2	v-myb avian myeloblastosis viral oncogene homolog-like 2
CEP55	Centrosomal protein 55kDa
MELK	Maternal embryonic leucine zipper kinase
ACTR3B	ARP3 actin-related protein 3 homolog B
NDC80	NDC80 kinetochore complex component
RRM2	ribonucleotide reductase M2
UBE2C	Ubiquitin-conjugating enzyme E2C
CENPF	Centromere protein F
PTTG1	Pituitary tumor-transforming 1
EXO1	exonuclease 1
ORC6L	Origin recognition complex, subunit 6
CCNE1	Cyclin E1
CCNB1	Cyclin B1
ANLN	Anillin, actin binding protein
KIF2C	Kinesin family member 2C

MYC	v-myc avian myelocytomatosis viral oncogene homolog
PHGDH	phosphoglycerate dehydrogenase
CDH3	Cadherin 3
MIA	Melanoma inhibitory activity
SFRP1	Secreted frizzled-related protein 1
SLC39A6	Solute carrier family 39 (zinc transporter), member 6
BAG1	BCL2-associated athanogene
MAPT	Microtubule-associated protein tau
CXXC5	CXXC finger protein 5
MLPH	melanophilin
BCL2	B-cell CLL/lymphoma 2
MDM2	MDM2 oncogene, E3 ubiquitin protein ligase
NAT1	N-acetyltransferase 1
BLVRA	biliverdin reductase A
MMP11	Matrix metallopeptidase 11
GPR160	G protein-coupled receptor 160
FGFR4	Fibroblast growth factor receptor 4
GRB7	Growth factor receptor-bound protein 7
TMEM45B	transmembrane protein 45B
ERBB2	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2
KRT17	Keratin 17
KRT14	Keratin 14
KRT5	Keratin 5
FOXA1	Forkhead box A1
FOXC1	Forkhead box C1

Appendix III: DEAD-box helicase DP103 defines metastatic potential of human breast cancers

Eun Myoung Shin,[1,2,3] Hui Sin Hay,[1,2,4] Moon Hee Lee,[5] Jen Nee Goh,[1,4] Tuan Zea Tan,[1] Yin Ping Sen,[6] See Wee Lim,[6] Einas M. Yousef,[7] Hooi Tin Ong,[8] Aye Aye Thike,[9] Xiangjun Kong,[10] Zhengsheng Wu,[10] Earnest Mendoz,[11] Wei Sun,[11] Manuel Salto-Tellez,[1,12,13]

Chwee Teck Lim,[11,14,15] Peter E. Lobie,[1,4,16] Yoon Pin Lim,[17] Celestial T. Yap,[3,18] Qi Zeng,[2,17] Gautam Sethi,[1,4] Martin B. Lee,[19]

Patrick Tan,[1,20,21] Boon Cher Goh,[1,18,22] Lance D. Miller,[23] Jean Paul Thiery,[1,2,17,18] Tao Zhu,[10] Louis Gaboury,[7] Puay Hoon Tan,[9]

Kam Man Hui,[8] George Wai-Cheong Yip,[6] Shigeki Miyamoto,[5] Alan Prem Kumar,[1,4,18,24,25] and Vinay Tergaonkar[2,17]

[1]Cancer Science Institute of Singapore, National University of Singapore, Singapore. [2]Institute of Molecular and Cellular Biology, A*STAR, Singapore. [3] Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [4]Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [5]McArdle Laboratory for Cancer Research, Department of Oncology, University of Wisconsin, Madison, Wisconsin, USA. [6]Department of Anatomy, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [7]Institute for Research in Immunology and Cancer (IRIC), University of Montreal, Montreal, Québec, Canada. [8]Division of Cellular and Molecular Research, Humphrey Oei Institute of Cancer Research, National Cancer Centre, Singapore. [9]Department of Pathology, Singapore General Hospital, Singapore. [10]Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, People's Republic of China. [11]Division of Bioengineering and Department of Mechanical Engineering,

National University of Singapore, Singapore. [12]Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [13]Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast, United Kingdom. [14]Mechanobiology Institute, National University of Singapore, Singapore. [15]NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore. [16]Liggins Institute, University of Auckland, Auckland, New Zealand. [17]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore. [18]National University Cancer Institute, Singapore. [19]Renal Center, National University Hospital, Singapore. [20]Cancer and Stem Cell Biology, Duke-NUS Graduate Medical School, Singapore. [21]Genome Institute of Singapore, A*STAR, Singapore. [22]Department of Haematology-Oncology, National University Hospital, Singapore. [23]Department of Cancer Biology, Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA. [24]School of Biomedical Sciences, Faculty of Health Sciences, Curtin University, Perth, Western Australia, Australia. [25]Department of Biological Sciences, University of North Texas, Denton, Texas, USA.

Published in The Journal of Clinical Investigation (JCI), in August 2014.

My contribution

I performed the IHC optimization and staining reaction.

I collected the data of Canada cohort of patients.

I did the scoring of the IHC reactions.

I prepared all the diagrams and pictures related to our experiments.

I carried out the statistical analysis related to our experiments.

I participated in the data analysis and interpretation regarding the IHC.

I participated in drafting the material and methods and results of the manuscript.

Abstract

Despite advancement in breast cancer treatment, 30% of patients with early breast cancers experience relapse with distant metastasis. It is a challenge to identify patients at risk for relapse; therefore, the identification of markers and therapeutic targets for metastatic breast cancers is imperative. Here, we identified DP103 as a biomarker and metastasis-driving oncogene in human breast cancers and determined that DP103 elevates matrix metallopeptidase 9 (MMP9) levels, which are associated with metastasis and invasion through activation of NF-κB. In turn, NF-κB signaling positively activated DP103 expression. Furthermore, DP103 enhanced TGF-β–activated kinase-1 (TAK1) phosphorylation of NF-κB–activating IκB kinase 2 (IKK2), leading to increased NF-κB activity. Reduction of DP103 expression in invasive breast cancer cells reduced phosphorylation of IKK2, abrogated NF-κB–mediated MMP9 expression, and impeded metastasis in a murine xenograft model. In breast cancer patient tissues, elevated levels of DP103 correlated with enhanced MMP9, reduced overall survival, and reduced survival after relapse. Together, these data indicate that a positive DP103/NF-κB feedback loop promotes constitutive NF-κB activation in invasive breast cancers and activation of this pathway is linked to cancer progression and the acquisition of chemotherapy resistance. Furthermore, our results suggest that DP103 has potential as a therapeutic target for breast cancer treatment.

Conflict of interest: Alan Prem Kumar is the principal inventor in a patent application based on invention disclosure entitled “Clinical Utility of DEAD-Box Helicase DP103 (DDX20) As a Prognostic and Predictive Biomarker for Drug Response in Cancer”

Introduction

Breast carcinoma is one of the most common malignancies in women around the world. Mortality from breast cancer is almost entirely the result of invasion and metastasis of neoplastic cells from the primary tumors to distant organ sites [135, 433, 510-515]; therefore, identifying genes involved in breast cancer metastasis is important. Since it is not possible to accurately predict the risk of metastasis in individual patients, at the present time, more than 80% of them receive adjuvant chemotherapy, but approximately 40% of the patients suffer relapse and ultimately die of metastatic disease. Cancer metastasis is a multistage process by which cancer spreads from the place at which it first arose as a primary tumor to distant locations in the body. These stages are interconnected through a series of adhesive interactions and invasive processes, including tumor angiogenesis, invasion, and colonization [437, 516-521]. Due to its heterogeneity and nature, mechanisms involving metastasis remain poorly understood. Degradation of the extracellular matrix (ECM) by matrix metalloproteinases (MMPs) [381, 522] is essential for almost every step in metastasis [453, 523]. It is now accepted that the strong and direct causal association of MMPs in breast cancers makes the inhibition of these enzymes a worthwhile strategy for impeding tumor metastasis. However, due to the multiplicity of MMPs and the widespread effects of their actions, inhibitors of MMPs have not been successful as anticancer agents. Indeed, there is a significant unmet need for identifying novel mechanisms to inhibit selective MMP functions in invading tumor cells.

DP103 belongs to the family of DExD/H-box proteins, named after the signature Asp-Glu-Ala-Asp/His motif within the helicase domain [524-530]. The 824-aa DEAD-box protein, DP103 (Gemin3, DDX20), was originally cloned and characterized as a component of the splicing machinery in concert with SMN, Sm, and other Gemin proteins [531-534]. DP103 is a transcriptional repressor for Egr2 in hind brain development [535] and forms a repressor complex with METS–PE-1 to silence transcription of Ets target genes involved in Ras-dependent macrophage proliferation and differentiation [530]. We have previously shown that DP103 transcriptionally represses the nuclear receptor SF-1 in a SUMO-dependent manner [527]. SUMOylation of SF-1 is mediated by PIASy, an E3 SUMO ligase, and this reaction is catalyzed in the presence of

DP103 as a cofactor. However, the role of DP103 as a cofactor in SUMOylation or other covalent modifications of cellular substrates and its involvement in cancer initiation or metastasis have never been explored.

The NF-κB family of transcription factors comprises 5 structurally related proteins that form homo- and heterodimers through their highly conserved DNA binding/dimerization Rel homology domain [536]. Binding of NF-κB to IκB proteins maintains NF-κB in an inactive state [537]. Activation of NF-κB in normal cells is inducible and is a tightly controlled event. Upon stimulation, IκBs are phosphorylated by the IκB kinase (IKK) complex (consisting of IKK1, IKK2, and NEMO proteins) [538-541]. IκB phosphorylation leads to its rapid proteolysis, thereby allowing NF-κB to function as a transcription factor. Temporally, IκBα is one of the first target genes that NF-κB activates very rapidly, and hence, this forms a feedback loop that shuts off the activity of NF-κB under normal circumstances. In many malignancies, including that of the breast, however, NF-κB is found to be constitutively activated [542-553]. Yet the molecular mechanism for such constitutive activation of NF-κB signaling in solid tumors is not clear. TGF-β–activated kinase-1 (TAK1), a member of the MAPK family, is a key regulator of signal transduction cascades, leading to the activation of the transcription factor NF-κB. Stimulation of cells with cytokines and microbial pathogens results in the activation of TAK1, which subsequently activates the IKK complex and MAPKs, culminating in the activation of NF-κB and AP-1, respectively [543, 544, 554-558]. TAK1 is recruited to the TNF-α receptor complex in an RIP-dependent manner following the stimulation of TNF-α receptor 1 and serves as a pivotal intermediate for IKK activation [543, 544, 548]. Apart from its critical role in immunity [559], wherein TAK1 is mostly activated by membrane or cytoplasmic cues, TAK1 is also essential for NF-κB activation in response to DNA damage, which is initiated from the nucleus [550, 560]. Although hypothesized, overexpression of molecules such as IKK2 and TAK1 as likely causes of constitutive NF-κB apparent in most malignancies has not been documented, despite sequencing of several cancer genomes.

In a search for markers of metastasis, we uncovered a molecule whose levels are upregulated in metastatic human breast cancers. Depletion of endogenous DP103 in invasive breast cancer cells led to decreased expression of matrix metallopeptidase 9

(MMP9) and impeded cell migration and invasion. Conversely, forced expression of DP103 in mammary carcinoma and non-malignant cells both in vitro and in vivo increased oncogenicity and the cells' capacity to invade. Mechanistic studies showed that DP103 is a critical cofactor for TAK1-mediated activation of IKK2, the key NF-κB–activating kinase in its activation loop. Hence, the observed effects of DP103 are mostly due its ability to activate NF-κB, an oncogenic transcription factor. Since a body of research indicates that TAK1-dependent NF-κB activation has a direct role in various aspects of cancer [543, 560], our identification of DP103 as the rate-limiting factor that mediates TAK1 signaling adds an important previously unidentified regulator that could be used as a drug target, since its levels specifically increase in cancer cells. Interestingly, NF-κB signaling also positively activates DP103 expression. Since levels of DP103 are limiting in normal cells, the existence of a tonic DP103–NF-κB– positive feedback loop in cancer cells could also explain how constitutive NF-κB activation is documented in most human cancers despite the absence of amplification of TAK1 or IKK2.

Results

Expression of DP103 is significantly upregulated in basal subtype human breast cancers.

Recently, a study showed increased expression of DP103 in a protein microarray done on tissues from mantle-cell lymphoma patients [561]. This prompted us to explore the role of DP103 in breast cancer. Tissue microarray (TMA) slides consisting of invasive ductal carcinoma (IDC) cases from 399 patients and normal nonmalignant ductal tissues from 61 women were obtained from the Department of Pathology, Singapore General Hospital. Clinicopathological features are shown in Supplemental Table 1A (supplemental material available online with this article; doi:10.1172/JCI73451DS1). Typical DP103 staining in normal versus tumor samples is shown in Figure 1, A and B, respectively. Based on DP103 staining, the cut-off value for low or high expression of DP103 was scored as 0 and 1+ versus 2+ and 3+. As a result of tissue loss during immunohistochemical processing, 330 cases of IDC and 38 cases of normal nonmalignant ductal tissues could be evaluated. As shown in Supplemental Table 2A, all

38 normal tissues showed low DP103 expression, while a significantly higher proportion (267 out of 330 patients) of specimens showed high DP103 expression in the tumor cores ($P < 0.001$). These findings were further extended using another cohort from the Montreal University Health Centre (Supplemental Figure 1, A and B, and Supplemental Table 2D). Clinico-pathological features are shown in Supplemental Table 1C. Taken together, these data sets reveal that expression of DP103 is significantly higher in tumors (across ethnic groups and regardless of the source of patient material).

To further delineate the expression profile of DP103 in various subtypes of breast cancer, 11 cohorts containing 1,325 breast tumors were collected and compiled from the NCBI's Gene Expression Omnibus (GEO) (see Methods). These 1,325 tumors were then classified using Single-Sample Gene Set Enrichment Analysis (ssGSEA) [562] and breast cancer subtype signature from Prat et al. [4]. As shown in Figure 1C, DP103 expression is significantly higher in basal subtype (Mann-Whitney test, $P = 4.88 \times 10^{-11}$). No significant difference in DP103 expression in claudin-low and luminal B were seen, while luminal A, ERBB2, and normal-like subtypes showed significantly lower expression (Mann-Whitney test, $P = 4.5 \times 10^{-5}$; $P = 0.0048$; $P = 0.0281$, respectively). Consistent findings were seen on a validation data set (GEO GSE3494) not included in the 11 meta-analysis cohorts, where basal subtype had significantly higher DP103 expression when compared against other subtypes (Mann-Whitney test, $P = 1.09 \times 10^{-4}$; see Supplemental Figure 1C). Since protein expression levels provide a more reliable quantification for function compared with mRNA quantification, as shown in Figure 1C and Supplemental Figure 1C, we then assessed protein expression of DP103 in the same 2 cohorts by immunohistochemistry (IHC). In agreement with our microarray results, the highest protein expression of DP103 correlated with the basal subtypes (Supplemental Figure 1, D–H; Supplemental Table 2B, Singapore cohort; Supplemental Figure 1, I–M; and Supplemental Table 2E, Canada cohort).

DP103 expression levels correlate with malignancy and with patient survival.

We then analyzed DP103 expression in breast tumors using a multi-institutional "microarray meta-analysis cohort" with a sample size of 669 primary breast cancer cases

and found levels of DP103 to be significantly elevated in poorly differentiated grade 3 tumors compared with those in well-differentiated grade 1 or 2 tumors (n = 669, P = 0.008) (Supplemental Figure 1N). We also validated the microarray data in Supplemental Figure 1N by analyzing DP103 protein expression by IHC. Corroborating our microarray results was the finding that DP103 protein is significantly higher in high-grade IDC compared with low-grade IDC (Supplemental Figure 2, A–D; Supplemental Table 2C, Singapore cohort; Supplemental Figure 2, E–G; and Supplemental Table 2F, Canada cohort).

Kaplan-Meier analysis using the cohort consisting of 399 patients (Supplemental Table 1A) revealed that high DP103 protein levels correlated with reduced survival. Patients with high DP103 expression (mean overall survival (OS), 132 months; n = 61) had significantly shorter survival (P = 0.010) compared with those with low DP103 expression (mean OS, 149 months) (Figure 1D). In addition, analysis of survival after relapse (SAR) revealed that patients with high DP103 expression (mean SAR, 63 months; n = 61) had significantly shorter survival (P = 0.009) compared with those with low DP103 expression (mean, 116 months) (Figure 1E). To investigate whether a meaningful functional correlation between DP103 expression levels and increasing tumor malignancy does exist, we determined basal DP103 expression in an isogenic cell xenograft–derived MCF10 breast cancer progression model system [563-565] that included the immortalized breast epithelial cell line MCF10A (10A1), the premalignant MCF10AT (AT1k), low-grade MCF10CA1h (CA1h), and high–metastatic grade MCF- 10CA1aCl.1 (CA1a) cell lines (Figure 1F). We found that DP103 protein (Figure 1G) and mRNA (Figure 1H) levels directly correlated with increasing invasive potential of these cell lines, with the highest levels of DP103 expression observed in the most metastatic CA1a cells (Figure 1, G and H).

DP103 expression correlates with metastasis gene signatures and breast cancer metastasis.

A recent US patent application on the use of compounds that block cancer metastasis showed DP103 (DDX20) as the only DEAD-box helicase member listed in the microarray data (publication number: US 2010/0004190 A1; ref. [566]), thereby

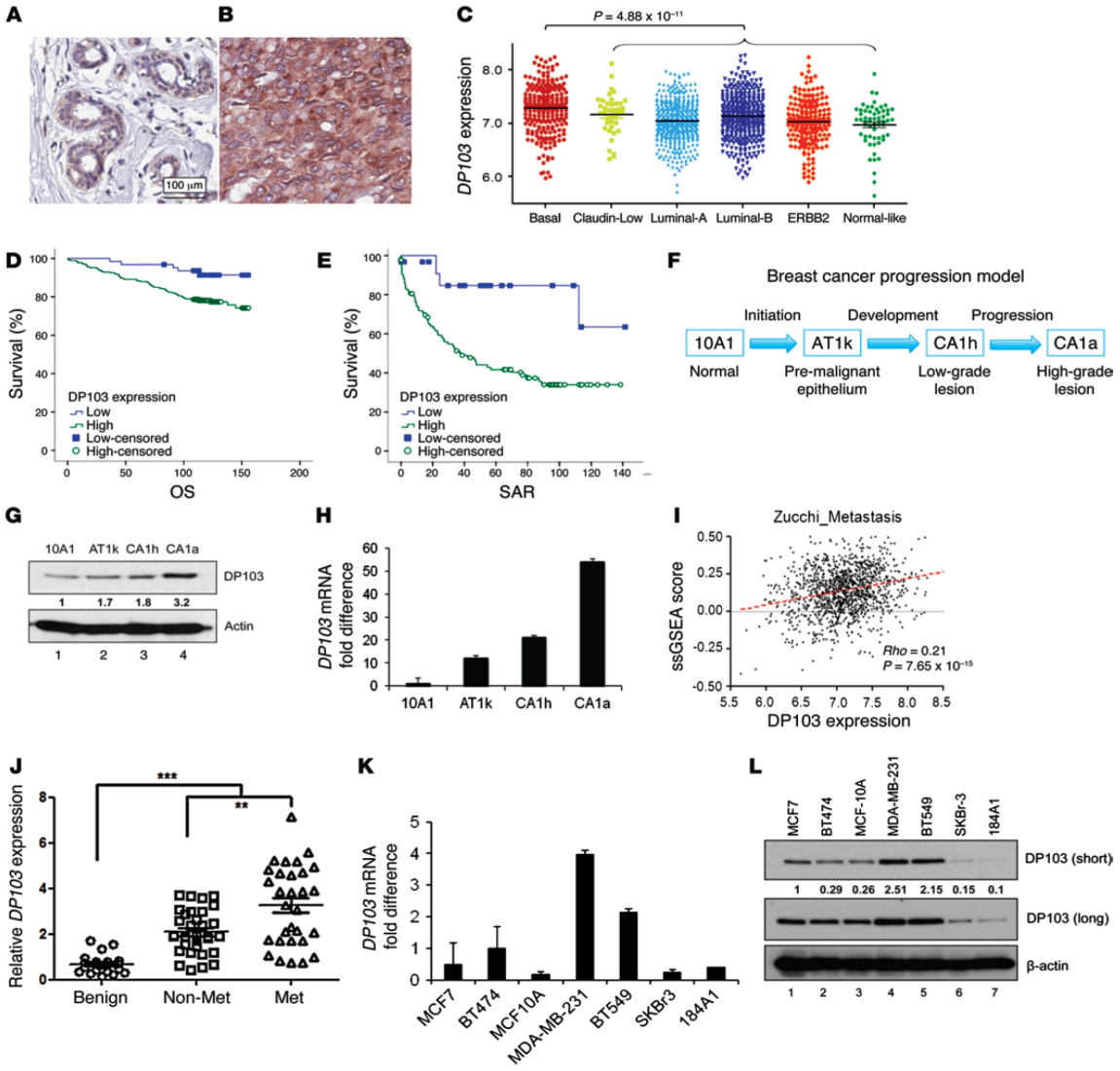

A
B
100 μm
C
P = 4.88 x 10⁻¹¹
DP103 expression
8.0
7.0
6.0
Basal
Claudin-Low
Luminal-A
Luminal-B
ERBB2
Normal-like
D
Survival (%)
DP103 expression
Low
High
Low-censored
High-censored
OS
E
Survival (%)
DP103 expression
Low
High
Low-censored
High-censored
SAR
F
Breast cancer progression model
Initiation
Development
Progression
10A1
AT1k
CA1h
CA1a
Normal
Pre-malignant epithelium
Low-grade lesion
High-grade lesion
G
10A1 AT1k CA1h CA1a
DP103
1 1.7 1.8 3.2
Actin
1 2 3 4
H
DP103 mRNA fold difference
10A1
AT1k
CA1h
CA1a
I
Zucchi_Metastasis
ssGSEA score
Rho = 0.21
P = 7.65 x 10⁻¹⁵
DP103 expression
J
Relative DP103 expression
***
**
Benign
Non-Met
Met
K
DP103 mRNA fold difference
MCF7
BT474
MCF10A
MDA-MB-231
BT549
SKBr3
184A1
L
MCF7
BT474
MCF-10A
MDA-MB-231
BT549
SKBr-3
184A1
DP103 (short)
1 0.29 0.26 2.51 2.15 0.15 0.1
DP103 (long)
β-actin
1 2 3 4 5 6 7

Figure 1. DP103 levels correlate with invasiveness and malignancy.

DP103 staining of (A) normal ductal tissue and (B) an IDC. (C) Gene expression value of DP103 (y axis) plotted for each breast cancer subtype, namely basal, claudin-low, luminal-A, luminal-B, ERBB2 (HER2+), and normal-like. (D) Kaplan-Meier curves showing DP103 expression in relation to patients' OS. Cases that have not experienced a positive event are censored at the date of last follow-up (small vertical lines on the line plots). (E) Kaplan-Meier curves showing DP103 expression in relation to SAR. Cases that have not experienced a positive event are censored at the date of last follow-up (small vertical lines on the line plots). (F) Breast cancer progression model showing isogenic cell lines with increasing invasive potential. (G) Western blotting with DP103 antibody in lysates from the isogenic cell lines (F). (H) qPCR with DP103 mRNA expression in RNA from the isogenic cell lines (F). (I) Gene expression of DP103 correlates with breast metastasis activity by Spearman correlation [65]. Red dotted line is curve fitted by linear regression. (J) Primary breast tissues from patients with benign disease, no lymph node metastases (Non-Met), and lymph node metastases (Met) collected and analyzed for DP103 mRNA expression **$P < 0.01$; ***$P < 0.001$. (K) RNA from breast cell lines and qPCR performed with DP103 primers. (L) Protein from breast cell lines extracted and levels of DP103 protein evaluated. Fold difference in protein expression indicated in G and L.

encouraging us to determine whether expression of DP103 correlates to the metastatic potential of tumor cells. Using breast cancer metastasis signature from Zucchi et al. [567], we employed ssGSEA to estimate the degree of cancer metastasis for breast tumors. Figure 1I shows a strong positive correlation in the meta-analysis cohort for breast cancer metastasis with DP103 expression (Spearman $\rho = 0.21$, $P = 7 \times 10^{-15}$), suggesting that DP103 may have a role in breast cancer metastasis. From this metastasis gene signature list [567], the association of the 2 most positively correlated metastasis genes, HMGB1 and H2AFZ, to expression of DP103 was then validated by quantitative PCR (qPCR) in a cohort of 63 breast cancer and 22 benign breast tissue samples derived from 85 patients who underwent surgery at the First Affiliated Hospital of Anhui Medical University, Hefei, China (Supplemental Figure 2, H and I). Clinicopathological status of this cohort is provided in Supplemental Table 1D. In this cohort, we then assessed whether expression of DP103 alone is predictive for lymph node metastasis. This analysis mirrored our gene expression analysis from above in that DP103 expression was lower in primary breast tissues that did not develop lymph node metastasis, while increased DP103 levels were observed in primary breast cancer tissues of patients who developed lymph node metastasis (Figure 1J). Benign breast tissue samples that were used as a reference demonstrated the least DP103 expression levels (Figure 1J). The role of DP103 in metastasis gained further credence by its listing in microarray data done on colorectal cancer (CRC) patients' tissues (n = 100), comparing gene expression profile in primary tumors with and without distant metastasis. When retrieved (ArrayExpress E-GEOD-18105), DP103 was found to be upregulated in CRC patients that developed distant metastasis ($n = 100$; $P < 1 \times 10^{-10}$). These results then encouraged us to screen a panel of breast cancer cell lines for DP103 expression. Interestingly, cell lines that are highly metastatic, such as MDA-MB-231 and BT549, display high levels of DP103 mRNA (Figure 1K) and protein (Figure 1L). On the other hand, non- invasive SKBr3, MCF7, and BT474 cells have much lower DP103 expression; while normal breast cells MCF10A and 184A1 have the lowest DP103 expression.

Suppression of DP103 decreases migratory ability of breast cancer cells.

To evaluate the functional consequence of DP103 expression in breast cancer

metastasis, we assessed cells' migratory behavior following RNAi-mediated knockdown of DP103 in metastatic MDA-MB-231 cells. To exclude off-target effects, we performed these assays using 2 independent siRNAs against DP103 (Supplemental Figure 3A) either singly or in combination. DP103 depletion led to decreased migration (by 40%) when siDP103#1 and siDP103#2 were used alone or (by 80%) when they were used together (siDP103 #1+2) (Figure 2, A and B) in a classical wound- healing assay. These observations were further confirmed by using a 2D nonwound migration assay (Supplemental Figure 3B). The 2D track plots of the cells are depicted in Supplemental Figure 3C, which shows reduction in cell motility, with the left panels showing cells transfected with control siRNA and the right panels showing cells transfected with siRNA against DP103. The top panels show the entire tracks made by the 50 cells, while the bottom panels show the ending points of each cell track (Supplemental Figure 3C). The data of individual cell tracks from this assay were then analyzed for various cell migration parameters, as shown in Supplemental Figure 3D. Accumulated distance is an indicator of motility of individual cells, whereas Euclidean distance gives an idea about the actual dispersion of the cells with respect to their starting positions. Confinement ratio is an indicator of directionality of the cells over time. In summary, we concur that suppression of DP103 affects the intracellular motility mechanisms in MDA-MB-231 cells but not the directionality of motion (Supplemental Figure 3D).

Suppression of DP103 decreases invasive ability of breast cancer cells.

Given that DP103 is upregulated in invasive breast cancer cells, we next explored effects of decreasing DP103 expression on tumor cell invasion by the use of Matrigel-coated Transwell chambers. In highly invasive MDA-MB-231 and BT549 cells, transient knockdown of DP103 resulted in about 50% less invasion through the Matrigel-coated inserts (Figure 2, C–E). The combined use of siDP103 resulted in about an 80% decrease in invading cells (Fig- ure 2, C–E). To mimic in vivo conditions as closely as possible, we then monitored luciferase-expressing MDA-MB-231 cells with and without suppression of DP103 expression using a 3D collagen gel assay. siDP103 reduced cells' invasiveness in 3D collagen hydrogel; notably, the spread of green tracks (ctsi) surpassed that of red tracks (siDP103) (Supplemental Figure 4A). Cell speed of the knock- down

cells was also averaged over each track and plotted into a histogram (Supplemental Figure 4B). As an additional indicator of invasive potential, MDA-MB-231 cells characteristically induce a branching morphogenesis. Interestingly, we observed a loss of this branching morphogenesis and reduction in cell spreading in cells depleted of DP103 expression (Supplemental Figure 4C).

Ectopic expression of DP103 enhances cell invasion.

Induction of tumor cell invasion is an important step in tumor metastasis and transition to a more malignant form of cancer at distant sites [6]. To further substantiate our findings, we then evaluated whether ectopic expression of DP103 would result in altered invasion by noncancer cells. Ectopic expression of DP103 alone had a dramatic effect on the phenotypic conversion of the MCF10A normal human breast epithelial cell line, increasing its invasion potential by 4-fold (Figure 2, F–H). In addition, forced overexpression of DP103 in MDA-MB-231 cells (Supplemental Figure 4, D–F) significantly promoted cells' invasive capacity. To validate these findings in vivo, we injected luciferase-expressing MDA-MB-231 cells stably transfected with either empty vector (MDA-MB-231-EV) or DP103 plasmid (MDA-MB-231-DP103) in the mammary fat pads of nude mice. Supplemental Table 3 compares the incidence of pulmonary and liver metastases at necropsy from primary tumors formed by the 2 breast cancer cell lines. Of 12 mice inoculated with MDA-MB-231-EV, only one was found to have lung and liver metastases by bioluminescence imaging. In contrast, 4 out of 11 mice injected with MDA-MB-231-DP103 cells developed distant metastasis to lung and liver (Supplemental Table 3). Figure 2I shows the ex vivo images of the lung and liver lobes from all mice that developed distant metastasis.

We validated differences in protein expression levels of DP103 in primary tumors of MDA-MB- 231-EV cells versus MDA- MB-231-DP103 cells by IHC (Figure 2J). In addition, we also quantified distal metastasis in lung and liver tissues by qPCR (human GAPDH gene expression normalized to mouse Gapdh gene expression). Our results show a 2-fold increase in distal metastasis to lung (Figure 2K) and liver (Figure 2L) in the MDA-MB-231-DP103 mouse group compared with the MDA-MB- 231-EV mouse group. Taken together, mice injected with MDA- MB-231-DP103 had higher incidence

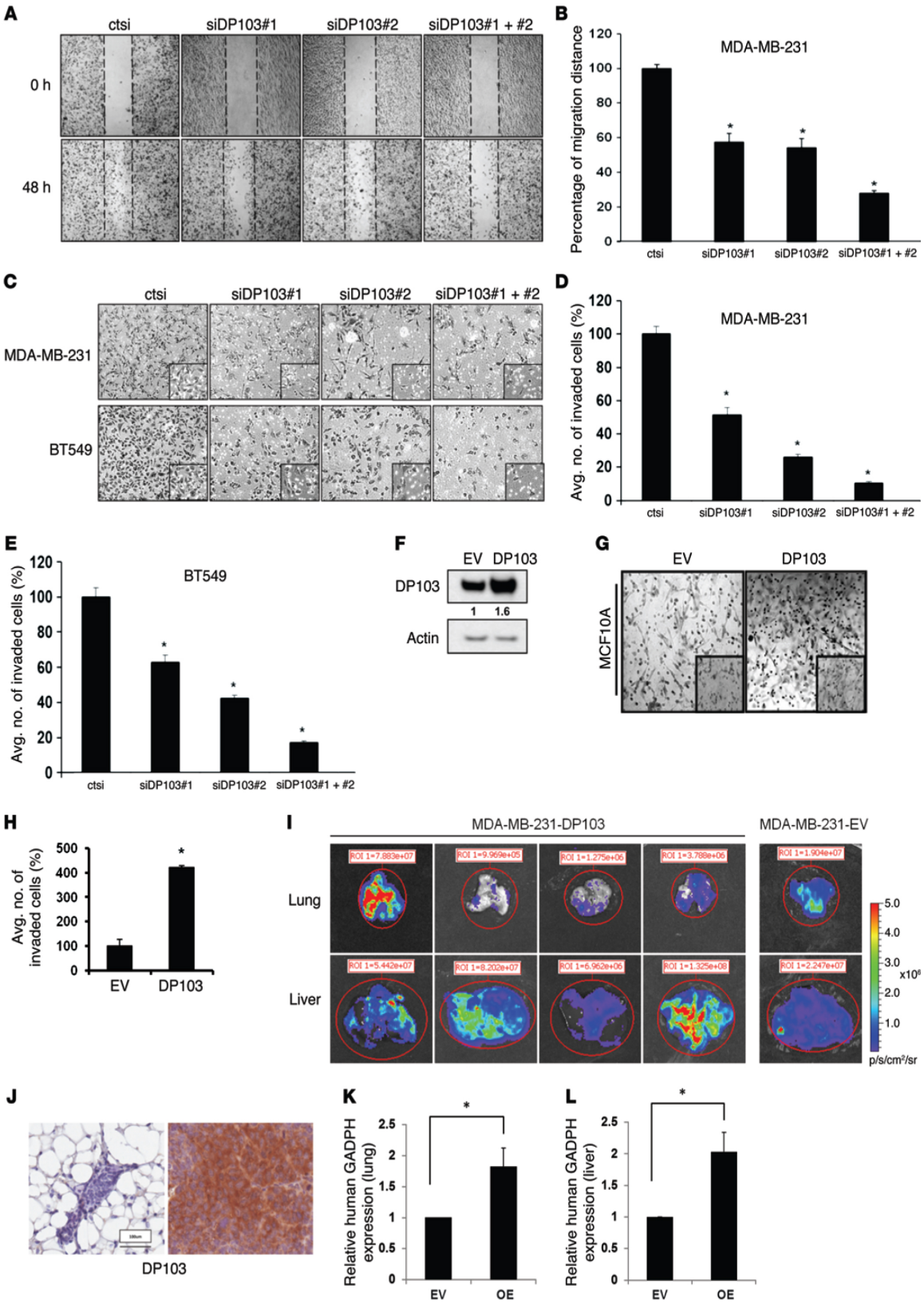

A
ctsi
siDP103#1
siDP103#2
siDP103#1 + #2
0 h
48 h
B
MDA-MB-231
Percentage of migration distance
C
MDA-MB-231
BT549
D
MDA-MB-231
Avg. no. of invaded cells (%)
E
BT549
Avg. no. of invaded cells (%)
F
EV
DP103
DP103
Actin
1
1.6
G
EV
DP103
MCF10A
H
Avg. no. of invaded cells (%)
EV
DP103
I
MDA-MB-231-DP103
MDA-MB-231-EV
Lung
Liver
p/s/cm²/sr
J
DP103
K
Relative human GADPH expression (lung)
EV
OE
L
Relative human GADPH expression (liver)
EV
OE

Figure 2. DP103 regulates invasive abilities of cancer cells.

(A) MDA- MB-231 cells transfected with control siRNA (ctsi) and 2 different siRNAs against DP103. Scratch wound healing assay is shown. (B) MDA-MB-231 cells represented in percentage of siDP103 or ctsi cell movement after scratching. *P < 0.05. (C) MDA-MB-231 and BT549 cells transfected with control siRNA (ctsi) and 2 different siRNAs against DP103 in Transwell assays. Inset shows representative photograph. (D) Quantification of MDA-MB-231 cells invaded through Transwell invasion chambers from C. *P < 0.05. (E) Quantification of BT549 cells invaded through the Transwell invasion chambers from C. *P < 0.05. (F) MCF10A cells transfected with control pcDNA3 vector (EV) and pcDNA3-FLAG-DP103 (DP103). Cell extracts immunoblotted with anti-DP103 and anti–β-actin antibodies. (G) MCF10A cells transfected as in F and analyzed in an invasion assay as in C. (H) MCF10A cells that invaded through the chambers quantified and represented. *P < 0.05. (I) Pulmonary and liver metastases from primary tumors evaluated by bioluminescence imaging. Color scale depicts photon flux (p/s) emitted from the organs. (J) Left: DP103 expression in primary tumor injected with empty vector–transfected MDA-MB-231 cells (MDA- MB-231-EV). Right: DP103 expression in mouse mammary fat pad injected with DP103-transfected cells (MDA-MB-231-DP103). (K) Distal metastasis to lung tissues quantified by mRNA levels of human GAPDH. *P = 0.05. (L) Metastasis to liver tissues quantified by mRNA levels of human GAPDH. *P = 0.05. Fold difference in protein expression indicated in F. Original magnification, ×10; ×40 (insets).

of lung and liver metastases than mice injected with MDA-MB-231-EV.

MMP9 levels predict survival and strongly correlate with DP103 levels in breast cancer.

Cell invasion requires the complex coregulation of cytoskeletal reorganization and cell motility as well as proteolysis and interaction with the ECM. Therefore, we measured expressions of several such genes in a "metastasis qPCR array" (Table 1). While ectopic expression of DP103 did not change the expression of genes such as NME1, PLAU, SERPINB5, or MTA1, it led to upregulation of several other genes, such as MMP1, TIMP1, and TIMP3. MMP9 mRNA levels were increased significantly (Table 1). Since high MMP9 expression is closely associated with poor prognosis in other cancer types [568-570], we stained tissues from the same cohorts used in Supplemental Table 1, A and D, for MMP9 protein expression. Similarly to expression of DP103, we observed high expression of MMP9 in breast tumor tissues compared with normal breast tissues (Figure 3, A and B, and Supplemental Table 4A, Singapore cohort; Supplemental Figure 5, A and B, and Supplemental Table 4B, Canada cohort).

Kaplan-Meier analysis showed that high MMP9 levels were correlated with higher tumor recurrence risk and reduced survival. Similarly to DP103, MMP9 status correlated with OS, with high expression (mean OS, 132 months; n = 69) showing significantly shorter survival duration ($P = 0.002$) compared with low MMP9 expression (mean OS, 150 months) (Figure 3C). Furthermore, cases with high MMP9 expression have shorter SAR (mean, 66 months; n = 69) and significantly shorter survival ($P = 0.028$) compared with those with low MMP9 expression (mean, 118 months) (Figure 3D). Having observed similar clinical significance for DP103 and MMP9, we then analyzed to determine whether there is a positive correlation between these 2 markers. Supplemental Table 4, C and D, indeed showed a strong positive clinical correlation between expression of DP103 and MMP9 ($P < 0.001$ in both cohorts). To further substantiate this finding, we next investigated whether a similar correlation between DP103 and MMP9 expression existed in the cohort from Supplemental Table 1D. Importantly, much like DP103, the highest levels of MMP9 were found in primary breast tumors that showed metastasis compared with nonmetastatic or benign tumors (Figure 3E). Importantly, the

correlation between DP103 and MMP9 expression was statistically significant ($P < 0.001$) (Supplemental Figure 5C) a trend that was also observed (Supplemental Figure 5D) in data obtained from the same microarray consisting of the 11 meta-analysis cohort used above (Figure 1C).

Table 1: Metastatic qPCR array

Gene	Gene name	Fold change (± SD)	*P* value
MTA2	Metastasis associated gene 2	1.173883 (±0.023156)	0.1418
NME1	Nucleoside diphosphate kinase A	0.935731 (±0.022723)	0.8644
PLAU	Urokinase-type plasminogen activator	1.089051 (±0.231132)	0.2914
MET	Hepatocyte growth factor receptor	1.458452 (±0.040862)	0.0036
PLAUR	Urokinase-type plasminogen activator receptor	1.154275 (±0.026637)	0.0499
MMP1	Matrix metalloproteinase 1	1.88218 (±0.165944)	0.0445
SERPINB5	Serpin peptidase inhibitor,clade B member 5	1.040501 (±0.335911)	0.5561
MMP2	Matrix metalloproteinase 2	0.80275 (±0.110911)	0.6882
SERPINE1	Serpin peptidase inhibitor, clade E member 1	0.845209 (±0.218096)	0.5754
MMP9	Matrix metalloproteinase 9	2.668643 (±0.000182)	0.0026
TIMP1	Metallopeptidase inhibitor 1	1.535293 (±0.062558)	0.0043
MTA1	Metastasis associated gene 1	1.225943 (±0.067582)	0.0387
TIMP3	Metallopeptidase inhibitor 3	1.659162 (±0.224292)	0.0372

Effects of ectopic expression of DP103 in breast cancer cells. RNA from BT549 stably transfected with either empty vector or DP103 was extracted and qPCR was performed with various primers. Results are represented as fold change in mRNA levels in BT549-DP103 cells relative to the control cells. Fold change values are representative of 3 independent mRNA replicates. Values less than 1 indicate decreased expression and more than 1 indicate increased expression of the specific mRNA. Values more than 2-fold changes are indicated in bold.

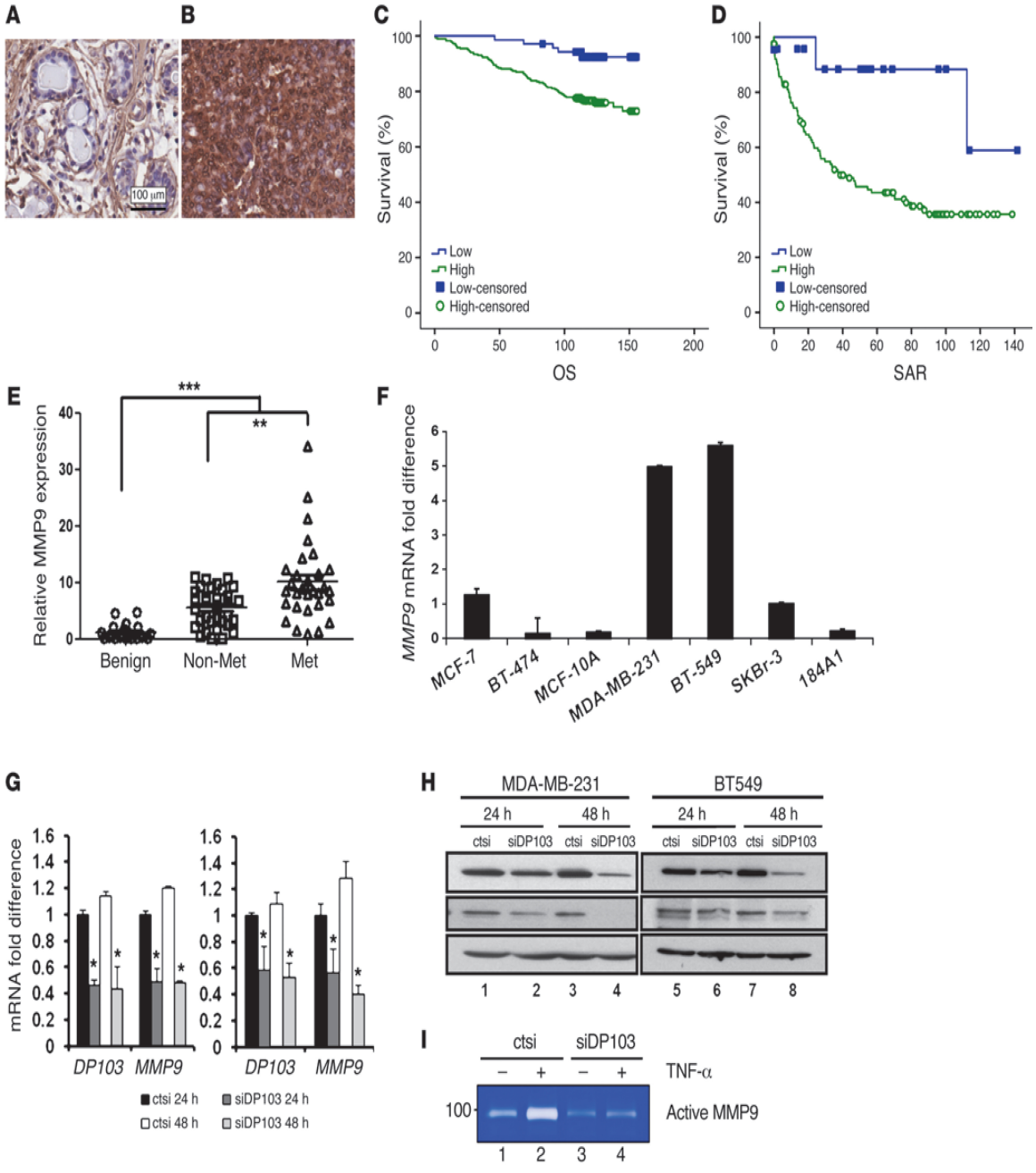

A
B
100 μm
C
Survival (%)
Low
High
Low-censored
High-censored
OS
D
Survival (%)
Low
High
Low-censored
High-censored
SAR
E
Relative MMP9 expression
***
**
Benign
Non-Met
Met
F
MMP9 mRNA fold difference
MCF-7
BT-474
MCF-10A
MDA-MB-231
BT-549
SKBr-3
184A1
G
mRNA fold difference
DP103
MMP9
DP103
MMP9
ctsi 24 h
siDP103 24 h
ctsi 48 h
siDP103 48 h
H
MDA-MB-231
BT549
24 h
48 h
ctsi
siDP103
1 2 3 4 5 6 7 8
I
ctsi
siDP103
TNF-α
100
Active MMP9
1 2 3 4

Figure 3. DP103 regulates MMP9 expression and function.

MMP9 staining of (A) normal ductal tissue and (B) IDC. (C) Kaplan-Meier curves of MMP9 expression in relation to OS. Cases that have not experienced a positive event are censored at the date of last follow-up (small vertical lines on the line plots). (D) Kaplan-Meier curves of MMP9 expression in relation to SAR. Cases that have not experienced a positive event are censored at the date of last follow-up (small vertical lines on the line plots). (E) Primary breast tissues from patients with benign diseases, no metastases, and displaying metastases analyzed for MMP9 mRNA expression. $^{**}P < 0.01$; $^{***}P < 0.001$. (F) Total RNA of different breast cell lines extracted and qPCR performed with MMP9 primers. (G) MDA-MB-231 and BT549 cells were transfected with control siRNA (ctsi) and siRNA against DP103. Total RNA analyzed for mRNA expression of DP103 and MMP9. (H) Expression of DP103 and MMP9 in cells used in G, evaluated using DP103 and MMP9 antibodies. (I) Suppression of DP103 in BT549 cells decreases MMP9 activity. Enzymatic activity of MMP9 was determined by gelatin zymography.

Suppression of DP103 expression decreases MMP9 gene expression.

We then screened for MMP9 expression in the panel of breast cell lines used in Figure 1K. We found that, mirroring the expression of DP103, metastatic breast cancer cell lines MDA-MB-231 and BT549 displayed the highest levels of endogenous MMP9 mRNA expressions, while the nonmetastatic and normal breast cell lines displayed relatively low levels (Figure 3F). In addition, we observed that MMP9 mRNA expression positively correlated with DP103 mRNA expression across a panel of cell lines (Pearson coefficient = 0.913) (Supplemental Figure 5E). Given the strong correlation observed between DP103 and MMP9 expression levels, we probed for a causal link between these enzymes. Transient knockdown of DP103 resulted in a significant decrease in MMP9 mRNA levels at 24 hours and 48 hours compared with cells transfected with a scrambled siRNA (ctsi) in MDA-MB-231, BT549 (Figure 3G), and Hs578t cells (Supplemental Figure 5F). Decrease in MMP9 mRNA corroborated well with an observed drop in MMP9 protein levels (Figure 3H and Supplemental Figure 5G) and enzyme activities (Figure 3I) without affecting cell viability (Supplemental Figure 5, H and I). Conversely, ectopic expression of DP103, much like what is seen in cancer tissues, resulted in increased MMP9 expression in both malignant MDA-MD-231 and nonmalignant MCF10A cells (Supplemental Figure 6A).

MMP9 mediates effects of DP103 in invasiveness.

In order to ascertain whether MMP9 is the primary molecular effector of the observed effects mediated by DP103, we evaluated invasion of MDA-MB-231 in a Transwell assay using a pan-MMP inhibitor (GM6001), a MMP2/9 inhibitor (SB-3CT), and a selective inhibitor of MMP9 (MMP9 inhibitor I). We found that inhibition by either GM6001 or SB-3CT prevented invasion of MDA-MB-231 cells through the Matrigel matrix by 40% (Supplemental Figure 6, B and C). More importantly, we found that specific inhibition of MMP9 activity alone was sufficient to significantly (~50%) impede the invasion of MDA-MB-231 cells either with basal or ectopic expression of DP103 (Supplemental Figure 6D), which suggests MMP9 is the key molecular effector mediating effects of DP103 seen in our experiments and in clinical settings. MMPs are known to play roles not only in metastatic progression, but also in tumor initiation and

formation [571], thereby suggesting that DP103-mediated regulation of these enzymes may have wide- spread implication in cancer initiation and progression. Indeed, although there was no difference in tumor size between control and DP103-overexpressing cells (Figure 2J and Supplemental Table 3), there were more tumors formed in DP103-expressing mice compared with control and more mice developed tumors in this group (data not shown).

DP103 regulates NF-κB–dependent gene expression.

Since an increase in MMP9 mRNA levels correlated with an increase in DP103 levels, it was likely that DP103 could be controlling the transcription of MMP9. Transcriptional regulation of MMP9 involves a relatively large repertoire of transcriptional factors that includes AP-1, AP-2, Ets, NF-κB, and SP-1, out of which AP-1 [572-574] and NF-κB [575, 576] are considered the major regulators. Given that depletion of DP103 expression in 3 invasive cell lines decreased MMP9 mRNA, we questioned whether this downregulation is due to inhibition of NF-κB or AP-1 activity. Suppression of DP103 led to downregulation of basal NF-κB activity, but had no effect on AP-1 activity (Figure 4A). Interestingly, DP103 levels were found to change in response to ecteinascidin 743 (ET-743), a genotoxic stress inducer [577]. Given that DP103 regulated MMP9 expression via activation of NF-κB, whose transcriptional activity is also regulated by genotoxic chemotherapeutic agents [550, 578-580], we then sought to investigate effects of DP103 in the presence of known chemotherapeutic agents that induce NF-κB, namely doxorubicin, etoposide (VP16), and camptothecin. Indeed, luciferase assays with NF-κB– and AP-1–dependent reporters showed that suppression of DP103 reduced drug-induced NF-κB reporter activity in both MDA-MB-231 and BT549 cells (Figure 4, B–D, and Supplemental Figure 7, A–C, respectively). Significantly, drug-induced AP-1 activity was not affected by DP103 levels (Supplemental Figure 7, D–F). EMSA analysis showed that siRNA-mediated reduction of DP103 reduced both basal and drug-induced NF-κB DNA–binding activity seen in both MDA-MB-231 and BT549 breast cancer cells (Figure 4E). AP-1 DNA–binding in both MDA-MB-231 and BT549 cells remained unaffected by DP103 levels and served as control (Figure 4E). Conversely, ectopic expression of DP103, much like what is seen in tumors, increased

NF-κB DNA–binding ability (Figure 4F). siRNA to DP103 also reduced nuclear accumulation of p65 while causing it to accumulate in the cytoplasm (Supplemental Figure 7G). We next evaluated whether DP103 could regulate other clinically relevant NF-κB target genes in response to stress. Indeed, siRNA to DP103 significantly reduced the levels of NF-κB target genes, namely ICAM-1, MMP9, and CXCR4 (Figure 4G).

Activation of NF-κB is known to protect cells from cell death induced by genotoxic agents. To evaluate whether reduction of NF-κB DNA binding (Figure 4, E and F, and Supplemental Figure 7G) and transcriptional activity (Figure 4, A–D, and Supplemental Figure 7, A–C) has functional significance, we treated MDA-MB-231 cells with doxorubicin or CPT and measured cell viability. Indeed, cells depleted of DP103 and hence with defective NF-κB activation were more sensitive to cell death induced by these genotoxic agents (Supplemental Figure 7, H and I). In addition, we found that silencing of DP103 led to decreased NF-κB binding to the DP103 promoter (Supplemental Figure 7J), suggesting that the DP103–NF-κB–positive feedback loop likely contributes to oncogenesis.

Activated NF-κB (p-p65-276) levels correlate with expression of DP103 and patient survival.

Since our results indicated that DP103, a key biomarker of cancer metastasis, functions via activation of NF-κB, we next explored whether there is a clinical value for NF-κB activation per se by evaluating expression levels of phospho-p65 in patient tissues. IHC performed on clinical specimens from the cohort shown in the Supplemental Table 1A revealed that, similar to expression of DP103, expression of phospho-p65 (Ser276), which represents the activated form of NF-κB, the competent form for chromatin remodeling, is significantly higher in breast tumor tissues compared with normal breast tissues ($P < 0.001$) (Figure 5, A and B, and Supplemental Table 5A). Analysis of phospho-p65 protein expression mirrored that of DP103, with the highest expression observed in basal subtypes compared with other subtypes ($P = 0.0141$) (Figure 5, C–G, and Supplemental Table 5B) and closely associated with high-grade tumors ($P < 0.001$) (Figure 5, H–K, and Supplemental Table 5C). Kaplan-Meier analysis on the same cohort revealed that high phospho-p65 (Ser276) expression (mean, 132 months; $n = 62$)

correlated with reduced OS, (P = 0.002) compared with low phospho- p65 expression (mean, 151 months) (Figure 5L). Furthermore high phospho-p65 (Ser276) expression (mean, 66 months; n = 62) predicted shorter SAR (P = 0.011) compared with low phospho-p65 expression (mean, 120 months) (Figure 5M). Most importantly, high expression of phospho-p65 correlated positively with high expression of DP103 in clinical specimens (P < 0.001) (Supplemental Table 5D).

DP103 regulates NF-κB in response to multiple stimuli.

NF-κB can regulate gene expression in response to multiple stimuli [543, 544, 581]. To test whether DP103 is an essential component of NF-κB signaling, we determined whether DP103 can also regulate NF-κB in response to a diverse array of well-known NF-κB–activating agents such as TNF-α, IL-1, LPS, CPT, and Dox. These agents initiate NF-κB activation via very distinct cellular-signaling players [544]. However all the distinct signaling to NF-κB initiated by these agents converges on the IKK complex. DP103 depletion impeded NF-κB DNA binding in response to all the tested stimuli (Figure 6A). DP103 also affected kinetics of NF-κB binding in response to membrane (TNF-α) and nuclear (CPT) activators of NF-κB (Figure 6, B and C). Furthermore, reduction of DP103 caused significant reduction in TNF-α–induced NF-κB–dependent transcription (Figure 6D).

Based on the finding that DP103 promoter contains κB sites (data not shown), we then evaluated whether DP103 itself is a target of NF-κB–activating stimuli. De novo DP103 protein was induced in response to both TNF-α and LPS (Figure 6E) and DP103 induction by LPS can be strongly and specifically inhibited by IKK2 inhibitor IV among other kinase inhibitors tested (Figure 6F), providing evidence that DP103 is an NF-κB target that could form part of a positive feedback loop. Interestingly, we identified 3 putative NF-κB– binding sites within the promoter of the DP103 gene (Figure 6G). By EMSA, we showed NF-κB binds to all 3 sites (data not shown). The specificity and identity of the NF-κB complexes could be ascertained using cold oligonucleotides and supershift with anti-p65 or anti-p50 antibodies, respectively (Figure 6H). Furthermore, EMSA with NF-κB consensus cold oligo-nucleotides abrogated NF-κB binding within the promoter of the DP103 gene (Figure 6I). In addition, we also showed a feedback loop

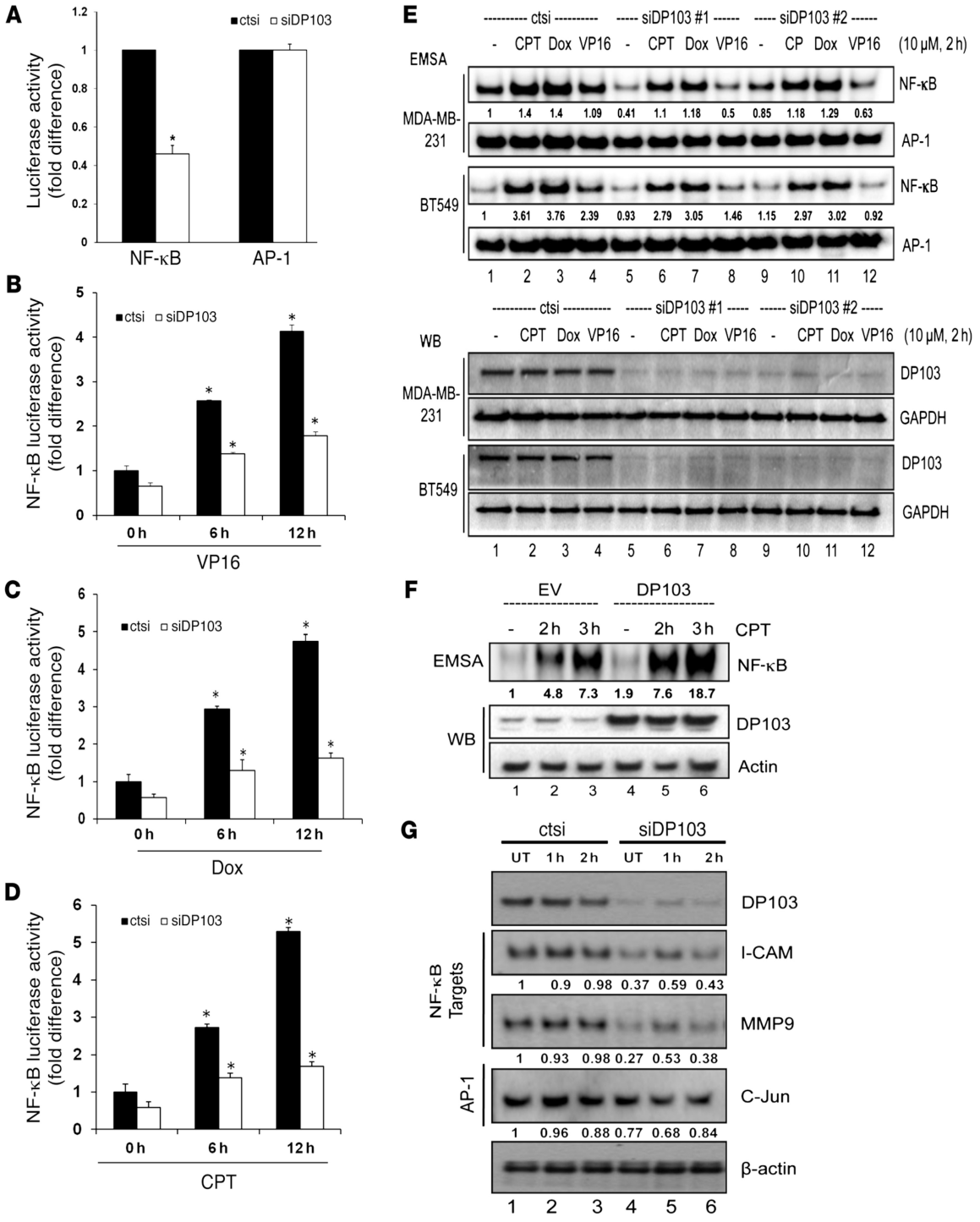

A
Luciferase activity (fold difference)
ctsi siDP103
NF-κB
AP-1
B
NF-κB luciferase activity (fold difference)
0 h
6 h
12 h
VP16
C
Dox
D
CPT
E
ctsi
siDP103 #1
siDP103 #2
- CPT Dox VP16 - CPT Dox VP16 - CP Dox VP16 (10 μM, 2 h)
EMSA
MDA-MB-231
BT549
NF-κB
AP-1
WB
DP103
GAPDH
F
EV
DP103
- 2 h 3 h - 2 h 3 h CPT
EMSA
WB
Actin
G
ctsi
siDP103
UT 1 h 2 h UT 1 h 2 h
NF-κB Targets
I-CAM
MMP9
AP-1
C-Jun
β-actin

Figure 4. DP103 is an essential regulator of NF-κB signaling.

(A) Control siRNA–treated (ctsi) and siDP103-treated MDA-MB-231 cells transfected with Renilla and luciferase reporter plasmid containing NF-κB or AP-1. Luciferase activity normalized and quantified. Results are average of 3 separate experiments. *P < 0.05. Control siRNA–treated (ctsi) and siDP103-treated MDA-MB-231 cells transfected with Renilla and luciferase reporter plasmid driven by NF-κB– binding sites. Cells were stimulated with (B) 10 μM VP16, (C) 25 μM doxorubicin, (D) or 10 μM CPT for 0, 6, and 12 hours and harvested for luciferase assays. Results are the average of 3 separate experiments. *P < 0.05. (E) MDA-MB-231 and BT549 cells transfected with either control siRNA (ctsi) or siDP103. Cells were left either untreated (–) or treated with CPT (10 μM), doxorubicin (10 μM), or VP16 (10 μM) for 2 hours. Protein extracts were analyzed by EMSA (top panel) and Western blotting using DP103 and GAPDH antibodies (bottom panel). (F) MDA-MB-231 cells infected either with lentiviral empty vector (lanes 1–3) or pBO- BI-DP103 (lanes 4–6) were either left untreated or treated with CPT (10 μM) for indicated times. Protein extracts analyzed by EMSA and Western blotting using anti-DP103 and anti–β-actin antibodies (top panel). (G) MDA-MB-231 cells transfected with either control siRNA (ctsi) or siRNA against DP103 and subjected to CPT (10 μM) stimulation at the indicated times. Cells were harvested and lysates evaluated by Western blotting for the indicated proteins. Fold difference in protein DNA binding indicated in E and F for EMSA and for protein expression changes indicated in G for Western blot.

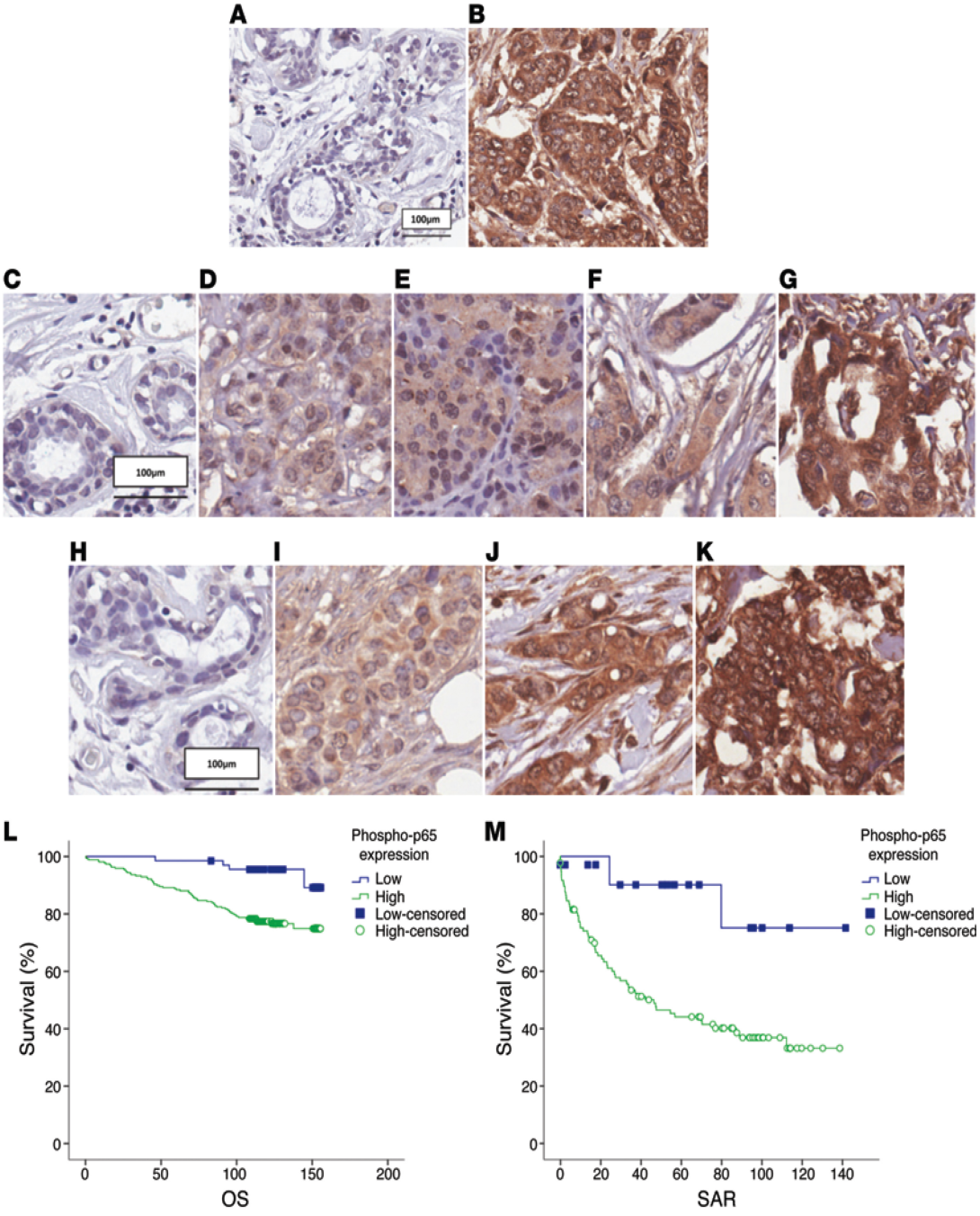

A
B
100μm
C
D
E
F
G
100μm
H
I
J
K
100μm
L
Phospho-p65 expression
Low
High
Low-censored
High-censored
Survival (%)
OS
M
Phospho-p65 expression
Low
High
Low-censored
High-censored
Survival (%)
SAR

Figure 5. Expression of active p65 (phospho-p65, Ser276) has a prognostic value. Staining of phospho-p65 (Ser276) in (A) normal ductal tissue and (B) IDC. Representative staining of phospho-p65 (Ser276) in (C) normal ductal tissue, (D) luminal A subtype, (E) luminal B subtype, (F) HER2 subtype, and (G) basal subtype. Phospho-p65 (Ser276) staining in (H) normal nonmalignant ductal tissue, (I) low-grade IDC, (J) IDC grade 2, and (K) IDC grade 3. Cell nuclei were counterstained with hematoxylin. (L) Kaplan-Meier curves showing phospho-p65 (Ser276) expression in relation to OS. Cases that have not experienced a positive event are censored at the date of last follow-up (small vertical lines on the line plots). (M) Kaplan-Meier curves showing phospho-p65 (Ser276) expression in relation to SAR. Cases that have not experienced a positive event are censored at the date of last follow-up (small vertical lines on the line plots).

following silencing of DP103 leading to decreased NF-κB binding to the DP103 promoter (Supplemental Figure 7G), suggesting that the DP103–NF-κB–positive feedback loop likely contributes to this oncogenic signaling arm in cancers.

Members of the helicase families, of which DP103 is a member, share several conserved motifs, including the Walker A and B motifs, which are involved in the binding of nucleoside triphosphates required for its helicase activity [524]. Changing the conserved GKT to GNT has been shown to reduce ATP binding in RNA helicases by 98% [582]. Additionally, the helicase-dead mutant (GNT) also retained its ability to induce invasion in MDA-MB-231 cells (Supplemental Figure 8, A–C), demonstrating that the helicase activity of DP103 is not required for its role in the metastasis of breast cancer cells. The helicase dead mutant of DP103 also interacts with TAK1 (Supplemental Figure 8D).

Given that a helicase dead DP103 functions in activating invasion and that DP103 can regulate signaling to NF-κB in response to a wide array of stimuli that converge only on the key TAK1-IKK complex, these results suggest that DP103 likely regulates a central complex in NF-κB activation such as the TAK1 or IKK complex independently of its role as a helicase. We tested kinase activity of the IKK complex on phosphorylation of IκBα with or without DP103 depletion in response to TNF-α as stimuli (Figure 6J). Phosphorylation of IκBα by the IKK complex was significantly impeded upon DP103 depletion. Furthermore, kinetics of endogenous IKK activation, as judged by phosphorylation of IKK1/2 in the activation loop at Ser180/Ser181 and its functional consequence, namely phosphorylation of IκBα (Ser32/Ser36), was markedly reduced when DP103 was depleted (Figure 6K), suggesting that DP103 is a key regulator of activating phosphorylations on the activation loop of IKKs and hence of IKK function in the physiological context.

DP103 enhances Tak1-mediated IKK2 phosphorylation and hence NF-κB activation.

Several mitogen-activated protein kinases, such as MEKK1, MEKK3, and TAK1, are suggested to regulate IKK activation. However, TAK1 is considered the immediate upstream activator of IKK and an essential component of both nuclear and receptor-

mediated NF-κB activation, which works by phosphorylating the Ser180/Ser181 in the activation loop of IKKs (39, 40, 44). We first investigated whether endogenous DP103 and TAK1 interact. Indeed, DP103 and TAK1 interacted in vivo, and this interaction was specific to the cytoplasm, since TAK1 is cytoplasmic (Figure 7A). Covalent modification of adaptor proteins by ubiquitin chains (termed as polyubiquitination) has been shown to be critical for activation of IKKs downstream of many NF-κB– activating stimuli. While there is much ongoing debate on the exact nature of the ubiquitin chains, such as in the K-63, linear, or K11 fashion, which is key for IKK activation, the ubiquitin-binding proteins TAB2/3 and NEMO are essential regulatory subunits of TAK1 and IKK complexes, respectively [583-585]. Indeed, deubiquitination enzymes A20 and CYLD, which remove these chains from adaptor proteins in NF-κB signaling, inhibit IKK and hence NF-κB activation. They are also known to be lost in inflammatory conditions and cancers, which explains the cause of hyperactivation of the pathway that characterizes these conditions [586].

Indeed, we could see TAK1 interacting proteins TAB2/3 interacting with DP103 under endogenous conditions (Supplemental Figure 9, A and B). To further verify whether this is a direct interaction, we purified GST-TAK1 and His-DP103 recombinant proteins (Supplemental Figure 9C) and tested their interaction in vitro (Figure 7B). Pull-down using Ni-NTA beads showed physical interaction between the 2 purified molecules (Figure 7B), suggesting that they indeed directly interact. To test the functionality of this interaction, we evaluated whether DP103 regulates the ability of TAK1 to phosphorylate IKK2 on its activation loop (40, 44). We purified the TAK1 phosphorylation site (residues 152–204) on IKK2 (GST-IKK2-WT) and a protein with mutations in the TAK1 phosphorylation domain (GST-IKK2-Mut) (Supplemental Figure 9D) and included these in a kinase assay with purified TAK1 and an increasing amount of DP103 (Figure 7C).

Increasing DP103 concentration caused an increase in TAK1-mediated phosphorylation (hence activation) specifically of IKK2-WT but not that of IKK2-Mut activation loops (Figure 7C). Importantly overexpression of DP103, much like what is seen in cancers, caused increased recruitment of IKK2 to TAK1 upon stimulation (Supplemental Figure 9, E and F). These results further explain the mechanism of constitutive IKK activation in cancers. Furthermore, given that the results in Figure 7C

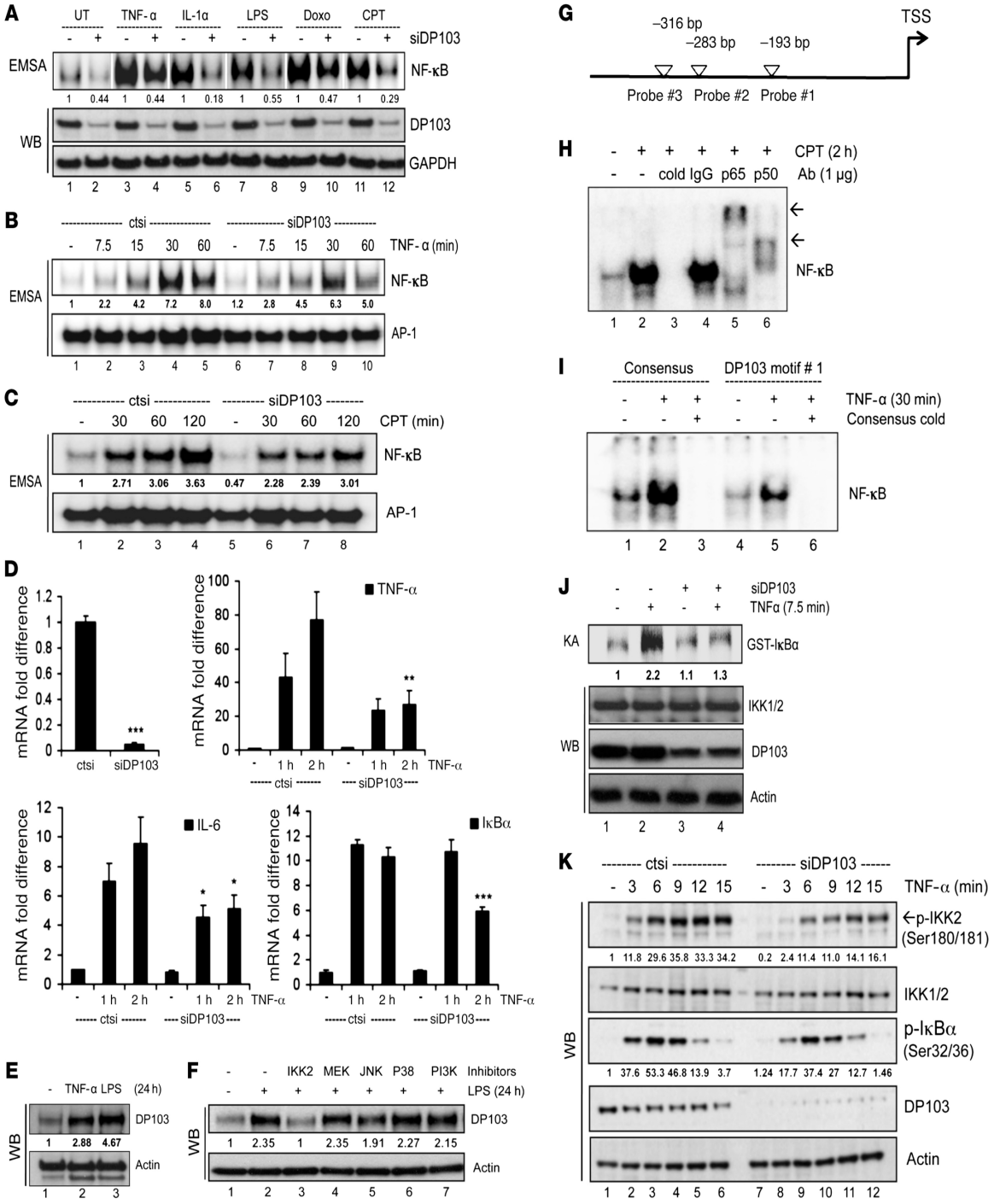

A
UT
TNF-α
IL-1α
LPS
Doxo
CPT
siDP103
EMSA
NF-κB
WB
DP103
GAPDH
B
ctsi
siDP103
TNF-α (min)
AP-1
C
CPT (min)
D
mRNA fold difference
ctsi
siDP103
TNF-α
IL-6
IκBα
E
F
IKK2
MEK
JNK
P38
PI3K
Inhibitors
LPS (24 h)
G
−316 bp
−283 bp
−193 bp
TSS
Probe #3
Probe #2
Probe #1
H
CPT (2 h)
cold
IgG
p65
p50
Ab (1 μg)
I
Consensus
DP103 motif # 1
TNF-α (30 min)
Consensus cold
J
TNFα (7.5 min)
KA
GST-IκBα
IKK1/2
DP103
Actin
K
TNF-α (min)
p-IKK2
(Ser180/181)
p-IκBα
(Ser32/36)

Figure 6. DP103 regulates NF-κB in response to multiple stimuli.

(A) MDA-MB-231 cells transfected with either control siRNA or siDP103. Cells untreated (UT) or treated as indicated. Extracts analyzed with EMSA (top) and Western blotting (bottom). White separating lines between different stimuli. (B and C) MDA-MB-231 cells transfected with control siRNA or siDP103. Lysates of cells untreated or treated with TNF-α (B) or CPT (C), tested by EMSA. (D) BT549 cells transfected as in C and stimulated with TNF-α. RNA analyzed for mRNA expression of NF-κB target genes. *P < 0.05; **P < 0.01; ***P < 0.001. (E) MCF10A cells treated with TNF-α or LPS. (F) MCF10A cells either untreated or treated with inhibitors for IKK2, MEK1/2, JNK, p38, or PI3K, then stimulated with LPS. (G) White arrowheads indicate putative NF-κB–binding sites on DP103 promoter. (H) Nuclear proteins from MDA-MB-231 cells either untreated or treated with CPT tested by EMSA. Unlabeled DP103 probe as cold or IgG, p65, or p50 was used. (I) Nuclear proteins from MDA-MB-231 cells untreated or treated with TNF-α tested by EMSA using NF-κB consensus (lanes 1–3) or DP103 probe (lanes 4–6). Unlabeled NF-κB consensus cold probe added to lanes 3 and 6. (J) Cells transfected either with siRNA control or siDP103. Lysates immunoprecipitated with NEMO antibody and kinase assay performed as previously described (36). (K) BT549 cells transfected either with control siRNA or siDP103 as indicated and immunoblotted. Fold difference in protein DNA binding indicated in EMSA for A–C and for protein expression changes in E, F, J, and K.

are derived using purified proteins, they suggest that DP103 can directly bind to TAK1 and function as a cofactor to enhance its kinase activity toward IKK2. This reaction is specific, since TAK1 activity toward the GST- IKK2-Mut protein remains unaltered. Indeed, if DP103 levels are limiting in normal cells and overexpression of DP103 as seen in cancers is the mechanistic basis of increased TAK1-mediated IKK and hence NF-κB activation apparent in cancers, this hypothesis could be tested. Next, we included the GST-IKK2-WT in a kinase assay where endogenous TAK1 was immunoprecipitated (Figure 7D). DP103 depletion showed reduced phosphorylation on Ser181 of GST-IKK2 by TAK1 (Figure 7D). Using MDA-MB-231 cells, we next evaluated whether DP103 levels influence the activation of the IKK kinase complex and hence NF-κB activation. Mere ectopic expression of DP103 increased activity of the IKK complex, as seen by phosphorylation of IκBα (Figure 7E). Conversely, depletion of endogenous DP103 expression led to reduced phosphorylation of IκBα (Figure 7E), reiterating that endogenous DP103 levels regulate TAK1 activity toward IKK2. Indeed, TAK1 depletion with siRNAs also showed downregulation of NF-κB target gene transcription (Figure 7F). Furthermore, much like siDP103 (Figure 3I), either TAK1 kinase activity inhibition or IKK2 inhibition or depletion with siRNAs against TAK1 abrogated TNF-α– induced MMP9 activation in MDA-MB-231 cels (Figure 7, G and H). In contrast, MMP2 was not affected, which is in line with the existing literature showing that MMP9 but not MMP2 is the bona fide NF-κB target. Finally, we tested functionality of these observations in MDA-MB-231 cells. Compared with control siRNA–treated cells, invasion of TAK1-depleted cells was greatly decreased in MDA-MB-231 cells (Figure 7I). We conclude that DP103-mediated regulation of TAK1's activity toward the major NF-κB kinase IKK2 regulates MMP9 expression and invasion.

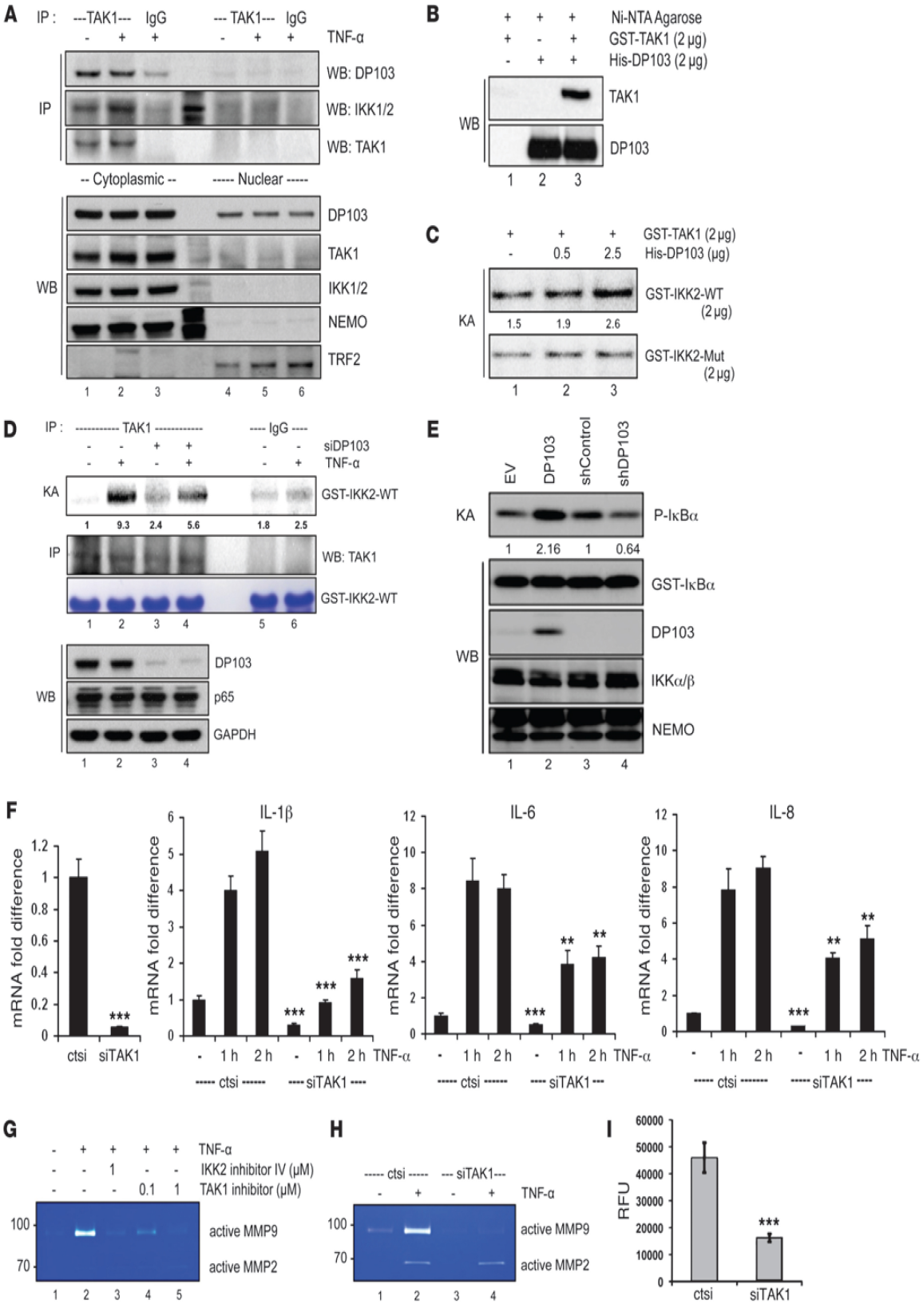

A
IP : ---TAK1--- IgG --- TAK1--- IgG
TNF-α
WB: DP103
WB: IKK1/2
WB: TAK1
IP
-- Cytoplasmic --
----- Nuclear -----
DP103
TAK1
IKK1/2
NEMO
TRF2
WB
1 2 3 4 5 6
B
Ni-NTA Agarose
GST-TAK1 (2 μg)
His-DP103 (2 μg)
TAK1
DP103
WB
1 2 3
C
GST-TAK1 (2 μg)
His-DP103 (μg)
0.5
2.5
GST-IKK2-WT (2 μg)
GST-IKK2-Mut (2 μg)
KA
1.5 1.9 2.6
1 2 3
D
IP : ---------- TAK1 ---------- ---- IgG ----
siDP103
TNF-α
KA
GST-IKK2-WT
1 9.3 2.4 5.6 1.8 2.5
IP
WB: TAK1
GST-IKK2-WT
1 2 3 4 5 6
DP103
p65
GAPDH
WB
1 2 3 4
E
EV
DP103
shControl
shDP103
KA
P-IκBα
1 2.16 1 0.64
GST-IκBα
DP103
IKKα/β
NEMO
WB
1 2 3 4
F
IL-1β
IL-6
IL-8
mRNA fold difference
ctsi siTAK1
- 1 h 2 h - 1 h 2 h TNF-α
----- ctsi ------ ---- siTAK1 ----
G
TNF-α
IKK2 inhibitor IV (μM)
TAK1 inhibitor (μM)
100
70
active MMP9
active MMP2
1 2 3 4 5
H
----- ctsi ----- --- siTAK1---
TNF-α
100
70
active MMP9
active MMP2
1 2 3 4
I
RFU
60000
50000
40000
30000
20000
10000
0
***
ctsi siTAK1

Figure 7. DP103 is a positive cofactor of TAK1-mediated IKK2 activation in MDA-MB-231 cells.

(A) Nuclear and cytoplasmic fractions immunoprecipitated with TAK1 and IgG control antibodies and immunoprecipitate material and lysates analyzed by immunoblotting with the indicated antibodies. (B) GST-TAK1 and His-DP103 incubated either separately or together. Immunoprecipitates analyzed by immunoblotting with the indicated antibodies. (C) Kinase assay performed using GST-TAK1 either with GST-IKK2-WT or GST-IKK2-Mut substrates with increasing input of His-DP103 protein. (D) Cells transfected either with control siRNA or siDP103 and stimulated with TNF-α. Lysates were immunoprecipitated using anti-TAK1 antibody, and kinase assay was performed using GST–IKK2 (amino acid residues 152–204) as substrate (top panel). (E) Cells transfected with DP103, siDP103, and respective control vectors. IKK complex immu- noprecipitated with anti-NEMO antibody. IKK activity was determined using phosphorylation of IκBα. (F) Cells transfected either with control siRNA or siTAK1 and stimulated with TNF-α. Total RNA analyzed for mRNA expression of NF-κB target genes. *P < 0.05; ***P < 0.001. (G) Cells treated with IKK2 inhibitor IV or TAK1 inhibitor (5Z)-7-oxozeaenol and stimulated with TNF-α. MMP9 activity evaluated with zymography. (H) Cells treated with control siRNA (ctsi) or siRNAs against TAK1 (siTAK1) and stimulated with TNF-α. MMP9 activity evaluated with zymography. (I) Cells transfected either with control siRNA or siRNAs against TAK1. Invaded cells through Matrigel detached and lysed in assay buffer were presented as relative fluorescence units (RFU). ***P < 0.001. Fold differ- ences in protein expression are indicated in C–E.

Discussion

Breast carcinoma is the most common malignancy in women, and its treatment is possible with early diagnosis. Despite high survival rates for early stage breast cancer, most such tumors often go unnoticed without regular mammograms. While effective targeted therapeutic modalities exist for women with hormone receptor–positive and HER2-positive disease, chemotherapy is the only systemic therapy available for women with triple-negative breast cancer. Increased heterogeneity of tumors in a larger group of patients poses a huge challenge in determining whether these drugs will show good efficacy in the OS and disease-free survival as well as the extent of the severity of side effects that these drugs will deliver. Strikingly, the number of patients who develop resistance to these therapies as well as patients who show relapse and metastasis is on the rise [587-589]. Metastasis is a multifactorial process requiring the concerted effort of many players. As such, it is a highly complex process and therefore unraveling key gene(s) whose expression levels could predict metastasis or tumorigenesis is of immense therapeutic value [518-520].

In this study, we uncovered DP103 as a prognostic marker and a therapeutic target for breast cancer metastasis. While members of the DEAD-box family, p68/p72 and Ddx3, have been linked to tumorigenesis, the potential role of DP103, a relatively new member [532], in cancer, if any, is yet to be uncovered. Using 2 independent cohorts, we found that a significant number of breast cancer patients display high levels of DP103 in their invasive breast tumor tissue, especially the basal subtype, compared with normal breast tissue. The prognostic and clinical significance of DP103 expression is highlighted by the observation that patients with high expression of DP103 have significantly shorter OS and SAR. Furthermore, DP103 levels were found to be upregulated in a statistically significant proportion of grade 3 breast cancers according to a multi-institutional mRNA microarray meta-analysis cohort of primary breast cancer; this was validated by IHC for DP103 protein expression in 2 independent cohorts. In vitro and in vivo analyses revealed that expression of DP103 strongly and positively correlates with the ability of breast cancer cells to metastasize. Remarkably, heightened expression of DP103 (as seen in cancers) in the normal epithelial breast cell line, MCF10A, which normally displays very low levels of DP103 and is unable to invade, was sufficient to confer a 4-fold increased

ability to these cells to invade. Overall, our study established that DP103 is not only a prognostic marker for breast cancer tumorigenesis and progression, but also a new predictive biomarker for cancer metastasis.

Since MMP9 expression has a clinical expression profile similar to that of DP103 and its expression is highly correlative to expression of DP103, both in clinical specimens and cell lines, we homed on the expression of the MMP family of proteins as a putative mechanism for observed effects of DP103. Our analysis further suggested that the ability of DP103 to regulate NF-κB but not AP1 activity may be crucial for its regulation of MMP9. Indeed, clinical expression of phospho-p65 (active NF-κB) mirrors the expression of DP103, both in terms of expression profile and survival outcomes, and most significantly, expression of phospho-p65 very significantly associates with expression of DP103. These results are significant because this study is the first, to our knowledge, to establish a role for DP103–NF-κB–MMP9 axis in breast metastasis.

How does DP103 regulate NF-κB–dependent transcription? Interestingly, DP103's helicase activity was not involved in the regulation of invasion. The only known helicase-independent activity of DP103 is its ability to affect PIASy-mediated SUMOylation of SF1 to act as a corepressor [527, 528]. However, the underlying molecular mechanism in the above 2 studies remains unknown. The known function of PIASy in the NF-κB–signaling system is limited to its ability to induce SUMOylation of NEMO in the DNA damage–signaling pathway [578, 590]. Indeed, DP103 could also affect PIASy SUMOylation, but this did not have a functional role in its regulation of MMP9 transcription (data not shown). Our results convincingly show that DP103-induced NF-κB and hence MMP9 activation is dependent on its ability to regulate TAK1-mediated activation of IKK via the phosphorylation of IKK activation loop. Knockdown and specific kinase inhibitors further prove that DP103-TAK1 complex is a potent activator of IKK2, the key NF-κB–activating kinase, and this induces MMP9 and hence invasion. Thus, identification of DP103 as a new limiting factor in TAK1-dependent NF-κB activation provides a significant mechanistic insight into the mechanism of constitutive NF-κB activation seen in tumors. It explains why, despite lack of overexpression of the TAK1 and IKK kinases, hyperactivation of NF-κB is apparent in the vast majority of cancers. It is due to the fact that DP103, levels of which increase in invasive or high-

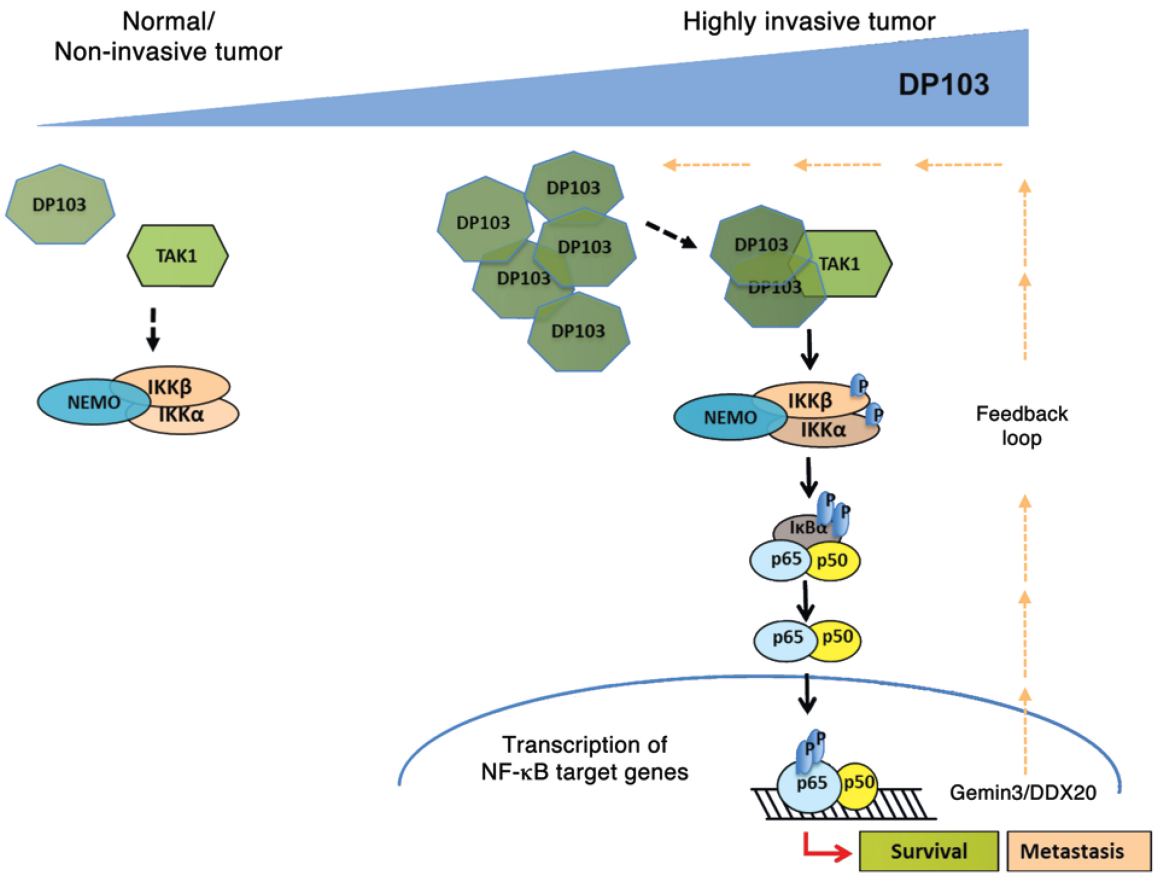


Figure 8. Schematic model based on our study showing a role for the RNA helicase DP103 in its ability to bind and stabilize TAK1 and thus activate NF-κB signaling in cancers.

grade cancer, provides a tonic activation of TAK1-mediated IKK2 and hence NF-κB activity. Since levels of DP103 are transcriptionally controlled by NF-κB itself in breast cancer cells, this feedback loop maintains constitutive activation of NF-κB and hence chronic inflammation, another hallmark of human cancers [591]. A model based on our studies is depicted in Figure 8. Since it is well accepted that constitutive NF-κB activity also imparts cancer cells the chemo- resistant phenotype [552, 553], DP103 could also be a therapeutic target in these cancers. Indeed, our data showed that reduction of NF-κB signaling by downregulating DP103 sensitized breast cancer cells to chemotherapy-induced cell death. DP103 on its own may not be a very attractive therapeutic target based on conventional targeting methods. Although small molecules and staple peptides have been designed to target "nondruggable" targets and even transcription factors such as BCL6, our finding that NF-κB itself can positively regulate DP103 expression provides us with an opportunity to break the positive feed-forward loop, the Achilles' heel of metastatic breast cancer, by blocking NF-κB activation. However, while blocking NF-κB by blocking IKK activity is successful in laboratory settings, this approach hasn't been successful in the clinic due to toxicity.

In summary, we describe a function of an RNA helicase DP103 that depends on its ability to bind and stabilize TAK1 and thus activate NF-κB signaling. Although constitutive activation of NF-κB is a well-documented phenomenon in cancer, increased levels of enzymes such as TAK1 that could maintain IKK2 and hence NF-κB in a constitutively active state are not seen in cancer cells. Instead, our current study uncovered that it is the increase in levels of DP103 that marks the switch from a nonmetastatic to a metastatic state in breast cancer cells and possibly other cancer cell types. In addition, we also elucidated a plausible DP103– NF-κB–positive feed-forward loop that could be involved in the maintenance of this oncogenic signaling arm in cancers. Thus, we suggest that DP103 is a novel biomarker as well as a worthwhile therapeutic drug target.

Materials and methods

Cell lines and reagents: MCF-7, MDA-MB-231, and BT549 human breast cancer cells were grown in RPMI with 10% FBS, 2 mM l-glutamine, and 0.05 mg/ml gentamicin. HEK293, BT474, and SKBr3 cells were grown in DMEM, DMEM/F12, and McCoy's 5A, respectively, with the same supplements as above. MCF10A and 184A1 cells were grown in MEGM with 10% FBS, 2 mM l-glutamine, bovine pituitary extract, and 0.05 mg/ml gentamicin. Xenograft-derived breast cancer cell lines (MCF10A1, MCF10AT1KCl.2, MCF10CA1h, and MCF10CA1a- Cl.1) were maintained in DMEM/F12 with 5% horse serum, 100 U/ml penicillin/streptomycin, 10 μg/ml insulin, 20 ng/ml EGF (Millipore), 0.5 μg/ml hydrocortisone, and 100 ng/ml cholera toxin. All cell lines except xenograft-derived were purchased from ATCC; media and supplements were from Hyclone unless otherwise indicated. Doxoru- bicin, VP16 (Etoposide), and camptothecin were purchased from Cal- biochem. All other chemicals and reagents were from Sigma-Aldrich unless otherwise indicated.

Other assays: Virus preparation and infection, transfections for siRNA and plasmids, Western blot analysis, immunoprecipitation, real-time PCR, GST-IκBα kinase, EMSA, cell viability, and luciferase assays were performed as described previously [540] and are described in detail in Supplemental Methods. Gel zymography [382] is described in detail in Supplemental Methods. Primer sequences are available in Supplemental Methods.

Metastatic qPCR array and microarray: Total RNA was extracted, and relative expressions of various genes were then analyzed using qPCR (ABI PRISM 7900; Applied Biosystems). Multiple gene markers distributed around the genome and 3 housekeeping genes were used for real-time PCR analysis using the SYBR GreenER qPCR SuperMix for ABI PRISM (Invitrogen), done in triplicate. Sequences of primers used were described previously [592, 593]. DP103 expression levels were analyzed in breast tumors using a multi-institutional microarray meta-analysis cohort totaling 759 primary breast cancer cases. The normalized microarray data set and associated clinical annotations were described previously [594]. Briefly, the microarray meta-analysis

cohort comprised data sets from 6 breast tumor-profiling studies. The data sets are accessible through MIAME-compliant public databases. Three data sets are housed in GEO [595] (GSE1456, ref. [596]; GSE6532, ref. [597]; and GSE9195, ref. [598]); 1 data set is housed in the NCI's caArray database (mille-00271; ref; [417]). All tumor samples were analyzed from frozen tissue collected at surgery and pro- filed on an Affymetrix U133 series microarray according to standard Affymetrix protocols [594]. Raw data (CEL files) were normalized by Robust Multichip Average (RMA) using the Bioconductor Affy Package (R) as previously described [599]. Batch effects between cohorts were corrected using the Partek Genomics Suite Batch Remover program. Two Affymetrix probe sets (223331_s_at and 224315_at) were designed to detect DP103 expression. The $\log_2$ signal intensities of these 2 probe sets were well correlated in the breast tumor meta- analysis cohort ($r = 0.65$, Pearson correlation).

Data preprocessing of Affymetrix microarray gene expression: Breast cancer data sets were downloaded from GEO (94). Microarray data on a U133Plus2 platform were utilized for analysis under the following GEO accession numbers: GSE12276 (n = 204), GSE19615 (n = 115), GSE21653 (n = 266), GSE23177 (n = 116), GSE23593 (n = 50), GSE26639 (n = 226), GSE3744 (n = 47), GSE5460 (n = 127), GSE5764 (n = 10), GSE6532 (n = 87), and GSE9195 (n = 77). For this study, we included all publicly available data set in the U133Plus2 platform at the time that the analysis was initiated in October 2011. RMA normalization was performed on each data set. The normalized data were compiled and subsequently standardized using ComBat [600] to remove batch effect. The standardized data yielded a data set of 1,325 breast cancer tumors and 20 normal breast tissue samples. The expression values of the DP103 gene were log averaged from DP103 probes in U133Plus2 platform: 223331_s_at and 224315_at. Apart from the 1,345 samples we compiled, we downloaded GEO GSE3494 (n = 252), data on both U133A and U133B platforms, and used it as validation data. GEO GSE3493 data were processed and normalized using the same method as described above.

ssGSEA: ssGSEA was originally described in a previous study [562].

Identification of breast cancer subtypes: Breast cancer subtype signature was obtained from Prat et al. (59). Subsequently, ssGSEA was computed based on the breast cancer subtype signature for each sample. Each sample was then assigned to be the subtype with the maximum ssGSEA score [4].

2D migrational assay: in vitro and 3D invasion assay, wound healing assay. 2D migrational assay, in vitro and 3D invasion assay, and wound healing assay were performed as per standard protocols and are described in Supplemental Methods.

Statistics: Statistical analyses were performed using the SPSS package (version 15.0 for Windows; SPSS Inc.), with significance set at the 5% level. Statistical significance evaluation by Mann-Whitney test and Spearman correlation test were computed using MATLAB. Dot plot was done using Graphpad Prism. Two-tailed Student's t tests were performed using the software Origin Pro, and results are represented as mean ± SD. Associations among DP103, MMP9, and phospho-p65 immunohistochemical staining data and clinico-pathological parameters were determined using Fisher's exact and Kendall tau tests. A P value below 0.05 was considered statistically significant. Survival curves were plotted using the Kaplan-Meier method and compared using the log-rank test.

Study approval: All animal studies were reviewed and approved by the SingHealth Institutional Animal Care and Use Committee, Sing- Health Office of Research, Singapore, and are described in Supplemental Methods. Ethics approval for the study using clinical material from Singapore was obtained from the Institutional Review Board, Singapore General Hospital. Protocol for the use of clinical material from China was approved by the Institutional Review Board of Anhui Medical University, while the use of clinical material from Canada was approved by the Institutional Review Board, Montreal University Health Centre, in 2005, with a renewal in 2010. Details on clinical materials and immunohistochemical procedures are described in Supplemental Experimental Procedures.

Acknowledgments

This work was supported by grants from the Singapore Ministry of Education Tier 2 (MOE2012-T2-2-139), the Academic Research Fund Tier 1 (R-184-000-228-112), and the Cancer Science Institute of Singapore, Experimental Therapeutics I Program (grant R-713-001-011-271) (to A.P. Kumar). A.P. Kumar is also supported by the John Nott Cancer Fellowship from Cancer Council, Western Australia. This work was also supported by the Singapore Ministry of Health's National Medical Research Council under its Individual Research Grant funding scheme (to G. Sethi); the National Key Scientific Program of China (2010CB912804, 2007CB914801, 2011CBA01103), and the CAS Visiting Professorship for Senior International Scientists (2010T2S03) (to T. Zhu). Core funding was received from A*STAR (to V. Tergaonkar) and NIH R01 CA77474 and GM083681 (to S. Miyamoto).

Supplemental information

SUPPLEMENTAL METHODS

Transfection, siRNA and plasmids

Two unique siRNAs targeting non-conserved coding regions of DP103 were used: #1: 5'- CCAGUGAUCCAAGUCUCAUAGCUUU-3'; #2: 5'- GCUGCCGCUUCUCAUU CAUAU UAUU-3' (Stealth RNAi^{TM}, Invitrogen, Life Technologies, Grand Island, NY). 5'-AGC UUC AUA AGG CGC AUG CTT (luciferase gene sequence inverted) was used as control siRNA (Qiagen, Valencia, CA). Full length-hFLAG-DP103 pcDNA3 was cloned from full length 2FLAG-hDP103, a kind gift from C. Glass (26). Helicase-dead mutant 2FLAG-hDP103-GNT pcDNA was derived from 2FLAG-hDP103 pcDNA by PCR-mutagenesis (Stratagene, La Jolla, CA) using sense primer 5'-TCTGGTACCGGGAATACCTGTGTGTTC-3' and antisense primer 5'-GAACACACAGGTATTCCCGGTACCAGA-3. The critical lysine residue needed to bind ATP (GKT) was mutated to an asparagines residue (GNT). For retroviral DP103 plasmid construction, a forward primer containing AgeI site and a reverse primer containing XhoI site was used to amplify the DP103 cDNA from Full length-hFLAG-DP103 pcDNA3. The amplified fragment was cloned into AgeI-XhoI cut pBobi plasmid, which contains a lentiviral backbone with the restriction sites placed immediately downstream of the Flag tag which is preceded by a CMV promoter. The primer sequences are: sense primer 5'-ATTAACCGGTATGGACTACAAGGAC-3' and antisense primer 5'-ATTACTCGA GTCACTGGTTACTATGCATC-3'.

Western Analysis

At the time of harvesting, the cell pellets were collected and completely lysed with RIPA lysis buffer (50mM Tris at pH 7.5, 150mM NaCl, 1% v/v NP-40, 1% v/v deoxycholic acid, 0.1% v/v SDS and 1mM EDTA) containing 1mM PMSF, leupeptin, pepstatin A and aprotinin before subjected to SDS-PAGE. The resolved proteins were then transferred onto nitrocellulose transfer membrane, blocked with 5% milk and

incubated with specific primary antibodies at 4°C overnight, followed by secondary antibodies before visualizing on X-ray films using enhanced chemiluminescence (Pierce, Rockford, IL). Human anti-mouse DP103 was purchased from BD Transduction Laboratories (San Diego, CA), human anti-rabbit MMP9, human anti-rabbit p- IκBα (Ser32/Ser36), human anti-rabbit ICAM, goat anti-mouse and goat anti-rabbit from Cell Signaling (Danvers, MA), human anti-rabbit TAK1 (H-579), TAB2 (H-300), TAB3 (H-128), TRF2 (N-20), c-Jun (H-79), NEMO (FL-419), IKK1/2 (H-470), p65 (C-20), p50 (H119), GAPDH (6C5), normal rabbit IgG and anti-mouse Myc (9E10) from Santa Cruz (CA, USA), anti-mouse Flag M2 from Sigma (St Louis, MO), anti-phospho-IKK1/2 (Ser180/181) from bioworld Technology (Louis Park, MN) anti-rat HA from Roche (Indianapolis, IN), and human anti-mouse SUMO-1 from Zymed (South San Francisco, CA).

RNA isolation, reverse transcription and realtime-PCR

Total RNA was extracted using TRIzol® reagent according to the manufacturer's instructions (Invitrogen, Life Technologies, Grand Island, NY). Reverse transcription (RT) was then carriedout. Each RT reaction contains 1μg of total RNA, 1x RT buffer, 5mM $MgCl_2$, 425μM each of dNTPs, 2μM random hexamers, 0.35U/μl RNase inhibitor, 1.1U/μl MultiScribe reverse transcriptase and made up to 10μl with sterile water. RT reaction was carried out at 42oC for 1h. The relative expressions of various genes were then analyzed using quantitative RT-PCR (ABI PRISM 7500, Applied Biosystems, Foster City, CA, USA) with 18S as an internal control. Primers and probes were purchased as kits from Applied Biosystems (Assays-on-Demand).

Invasion assay

In vitro invasion assay was performed using BD Bio-Coat Matrigel invasion assay system (BD Biosciences, San Jose, CA) according to the manufacturer's instructions. Briefly, cells were trypsinized 48h post-transfection, $2x10^5$ cells suspended in serum-free medium and seeded into the Matrigel transwell chambers consisting of polycarbonate membranes with 8-μm pores. After incubation for 24h, the upper surface of transwell chambers was wiped off with a cotton swab and invading cells were fixed and stained

with crystal violet solution. The invading cell numbers were counted in five randomly selected microscope fields (x200) and their averages were converted to percentage with the control setup taken to be at 100%.

Immunoprecipitation

A total of $3x10^6$ cells were treated as indicated. Cell pellets were lysed in lysis buffer (50mM Tris (pH 7.4), 150mM NaCl, 0.5% NP-40, 0.5mM PMSF, 20mM β-glycerol phosphate, 1mM sodium orthovanadate, 1μg/ml leupeptin, 1μg/ml aprotinin, 50mM sodium fluoride). Pre-cleared 0.5mg of proteins were incubated with Protein A Sepharose, CL-4B (GE Healthcare Life Science, Pittsburgh, PA) and indicated antibodies (1μg) overnight. Beads were washed four times with (200mM Tris (pH 8), 100mM NaCl, 0.5% NP-40, 2mM DTT, 0.5mM PMSF, 1mM sodium orthovanadate, 1μg/ml leupeptin, 1μg/ml aprotinin). The samples were boiled in 4x SDS loading buffer, and proteins were separated by SDS-PAGE and subjected to Western blotting with the indicated antibodies.

Wound healing assay

MDA-MB-231 and BT549 cells were treated according to experimental design. Before plating the cells, two parallel lines were drawn at the underside of the well. These lines served as fiducial marks for the wound areas to be analyzed. In preparation for marking the wound, the cells should be fully confluent. The growth medium was aspirated and replaced by calcium-free PBS to prevent killing of cells at the edge of the wound by exposure to high calcium concentrations. Two parallel scratch wounds were made perpendicular to the marker lines with a 1000μl blue tip. The medium was then changed to complete media. After incubation for 48h, the wounds are observed using bright field microscopy and multiple images were taken at areas flanking the intersections of the wound and the marker lines at the start and end of the experiment. Three measurements of the gap distance between the wound were taken at the start and end of the experiment from the images, and their averages are converted to percentages to depict percentage change with the control setup taken to be at 100%.

Electrophoretic mobility shift assay

Cell pellets were lysed in Totex buffer (20mM HEPES (pH 7.7), 350mM NaCl, 20% glycerol, 1% NP-40, 1mM $MgCl_2$, 0.5mM EDTA, 0.1mM EGTA, 0.5mM DTT) containing protease inhibitors, spun down at 14,000 x *g* for 10min at 4°C and subjected to EMSA as described previously (114). The supernatant collected was divided into three separate samples for analysis with double-stranded NFκB, AP-1 and Oct-1 (γ-^{32}P) radiolabeled probes. NF- κB: 5'-TCA ACA GAG GGG ACT TTC CGA GAG GCC-3' (115) AP-1: 5'- CGC TTG ATG ACT CAG CGG GAA-3' (116)
Oct-1: 5'- TGT CGA ATG CAA ATC ACT AGA A-3' (117)
The sequences of (γ-^{32}P) radiolabeled probes used to confirm NF-κB binding sites within the promoter of DP103 are:
probe#1: 5'-TCTCCTCCCTCTTGGGGCTTTCCT-3',
probe#2: 5'-AGAGGCGGGGCGGTGCCCCCACCG-3' and
probe#3: 5'-CACGGCTGGGCGGCTCCGCCCAG-3'.

The DNA probe was radiolabeled by incubation at 37°C for 30min with the following: 5x forwarding buffer, ^{32}P-γ-ATP and T4 kinase (10unit/μl). The labeled probe was then spun at 3000 rpm for 1min to get rid of the buffer, followed by loading into the G50 micro column (Pharmacia Biotech, UK) and spun for 2min for purification. EMSA was performed in a reaction mixture containing 2x reaction buffer (50% glycerol, 1M Hepes pH 7.9, 1M Tris-HCl pH 8.0, 0.5M EDTA pH 8.0, 100mM DTT), 0.5μg/μl poly (dI-dC), protein samples and 1μg/μl BSA and kept on ice for 10min before adding DNA probes. The mixture was then kept at room temperature for another 20min before loading into a non-denaturing polyacrylamide gel and run for 1h. The gel was then dried and analyzed with a PharosFX Plus system (BioRad, Hercules, CA).

Gel zymography

Equal serum-free growth media collected from tested samples as indicated were subject to Electrophoresis using 10% denaturing polyacrylamide gels containing 0.1% gelatin. Following electrophoresis, the gels were re-natured with 50mM Tris-HCl buffer (pH 7.5) containing 2.5% Triton X-100 for 30min, and incubated at 37°C for 16h in a

buffer composed of 0.15M NaCl, 10mM $CaCl_2$ and 50mM Tris-HCl (pH 7.5). The gels were stained with 0.5% Coomassie blue in 5% methanol and 10% acetic acid in dH_2O, and destained with 10% methanol and 5% acetic acid.

Luciferase assay

$1.25x10^5$cells/well was plated in 12-well plates. Experiments were set up as described. The cells were then transfected with luciferase reporter plasmid containing 3x NFκB or AP-1 binding sites together with *Renilla* plasmid (Clonetech, Palo Alto, CA). At time of harvest, the promoter activity was assessed with a dual-luciferase assay kit (Promega, Madison, WI). Briefly, feeding medium was removed from the wells, washed once with 1x PBS, and lysed with ice-cold 100µl of reporter lysis buffer. Ten microlitres of cell lysate was then added to 50µl of luciferase substrate solution, following which 50µl of stop & glow buffer was added for *Renilla* reading. Bioluminescence generated was measured using a Sirius luminometer (Berthold Detection System, Pforzheim, Germany). The luminescence readings obtained were normalized to the protein concentration of the corresponding cell lysate and presented as fold difference with reference to the control setup.

Cell viability

MDA-MB-231 cells were plated in 12-well plates. Experiments were set up as described. At the end of drug treatment or transfection, medium was removed from wells. The cells were then washed once with 1x PBS. This was followed by incubation with 0.5ml of crystal violet solution (0.75% crystal violet, 50% ethanol, 1.75% formaldehyde, 0.25% NaCl) for ten minutes. Excess crystal violet solution was carefully washed away with distilled water for several times and the wells were left to air-dry. The remaining crystals were dissolved in a 1% SDS in 1x PBS solution and its absorbance read at 595nm and converted to percentages, with control setups at 100%.

Mammary fat pad spontaneous metastasis model

Ten-week old female Balb/c nude mice (Animal Resource Centre, Western Australia) were anaesthetized prior to surgery and a 5mm incision was made in the skin

to expose the abdominal mammary fat pad (m.f.p.). Two million cells in 0.03ml were injected into the tissue through a 27- gauge needle. Tumor growth was monitored weekly by bioluminescence imaging using the IVISTM camera system (Xenogen, Alameda, CA, USA). For *ex vivo* imaging, 150mg/kg D- luciferin (Xenogen) was injected into the mice just before necropsy. Tissues of interest were excised, placed into tissue culture dish and imaged for 1min. Regions of interest from displayed images were quantified as photons per second (p/s) using Living Image Software (Xenogen).

2D migrational assays

Glass cylinders (Bioptechs, Butler, PA) of 6mm inner diameter were placed vertically on tissue culture dish. About 20000 cells were seeded inside the glass cylinders and incubated at 37°C in a humidified atmosphere at 5% CO2 for 24h. The cylinders were then carefully lifted from the dish to reveal an undisturbed circular monolayer patch of cells, which were then washed thrice with 1x PBS to remove dead cell debris and refilled with 2ml of complete medium. Live video monitoring assays of the migrating cells at the edge of the monolayer were performed using phase contrast microscopy (Biostation IM, Nikon). Rectangular fields of view with pixel resolution 1280x960 were chosen from the monolayer edge using the proprietary software and videos were recorded for 24 hours with 10min intervals in between the frames (total 145 frames).

Fifty cells from 6-8 rows of the leading edge of the monolayer were manually tracked using the open source software Image J. Monolayer edge distances (MED) were measured as the average of the displacements (n=5 per field of view measured for 4-5 videos per experiment) between the initial and final positions of the monolayer edge. 2D track plots and plot related measurements were performed using the Chemotaxis tool plugin (Integrated BioDiagnostics).

In vitro and 3D invasion assay

In vitro invasion assay was performed using BD Bio-Coat Matrigel invasion assay system (BD Biosciences, San Jose, CA) according to the manufacturer's instructions. For 3D invasion assay, NutragenTM collagen solution (Inamed Biomaterials, Fremont, CA, US) was mixed with NaOH, 10x PBS, and MDA-MB-231-GFP cells suspended in

serum-free DMEM on ice. The final solution contained 4mg/ml of collagen-I and around 200,000/ml cells. 30-60min incubation under 37°C humidified chamber led to the self-assembly of a piece of semi-spherical cell-seeded collagen gel in the central well of a glass-bottomed dish. Complete cell culture media were immediately supplied to the gel to support cell growth. After 40h, these GFP-expressing cells were imaged in 3D collagen hydrogel using confocal fluorescence microscopy (Nikon TE2000- EZ C1 system). Ten hours of time-lapsed confocal imaging was then carried out in an environment chamber maintaining 37°C and 5% CO2 atmosphere, which was completed within 70h post siRNA transfection. Quantitative image analysis was achieved using 3D reconstruction and cell tracking function provided by Imaris (Bitplane, Zurich, Switzerland).

GST-IκBα Kinase Assay

MDA-MB-231 cells were transfected with DP103 for overexpression or siDP103 for knock- down. After 36 h, the cells were lysed with IP buffer and the kinase complex was prepared by IP using anti-NEMO antibody. GST-IκBα/1-66 was purified using glutathione-agarose column (Thermo Sci., MA) and used as the substrate. The reaction was performed with mixture of kinase complex, 0.5μg of GST-IκBα/1-66, and ATP (P^{32}), in kinase buffer (20 mM Hepes (pH 7.7), 2mM $MgCl_2$, 2mM $MnCl_2$, 1mM DTT, 10μM ATP + inhibitors (0.5 mM PMSF, 10mM beta- glycerol phosphate (BGP), 300μM sodium othovanadate, 1μg/ml leupeptin, 1μg/ml aprotin, 10mM sodium fluoride, 10mM p-nitrophenyl phosohate) at 37°C for 1h. The samples were subjected to 12% SDS-PAGE and visualized by autoradiography. The purified IKK comlex and GST- IκBα/1-66 were analysed by western blotting with proper antibodies.

Clinical materials

Table S1A&B – Singapore Cohort

Tissue microarray (TMA) slides consisting of invasive ductal carcinoma (IDC) cases from 399 patients and normal non-malignant ductal tissues from 61 women were obtained from the Department of Pathology, Singapore General Hospital. As a result of tissue loss during immunohistochemical processing, the following number of cases (invasive ductal carcinoma and normal non-malignant ductal tissues) were available for

evaluation: DP103 (Proteintech, Chicago, IL, USA (catalogue 11324-1-AP)) (330 and 38 respectively); phospho-p65 (S276) antibody (Abcam, Cambridge, MA, USA (catalogue ab30623)) (338 and 49 respectively); MMP9 (Proteintech, Chicago, IL, USA (catalogue 10375-2-AP)) (357 and 53 respectively). Clinicopathological features were recorded for statistical analyses, including age of patient, ethnic group, tumor type, and histological grade of tumor. Patient survival and tumor recurrence data were available for 329 cancer patients. The period of follow-up ranged from 0 to 156 months. Deaths (defined as being resultant from the cancer itself) occurred in 21.1% of patients, with the mean and median overall survival (OS) periods being 112 and 117 months respectively. Among patients with tumor recurrence, 20.1% of them were dead at the end of the study period, with a mean survival after recurrence (SAR) period of 17 months and median SAR of 0 month. OS was defined as the time from diagnosis to death. SAR was defined as the survival duration from the first recurrence to death. Cases that did not reach the defined end-points of interest were censored at the date of last follow-up.

Table S1C – Canada Cohort

This study was performed on 190 archived formalin fixed paraffin embedded (FFPE) blocks containing tissues obtained from female patients. All samples were obtained from Centre hospitalier de l'Université de Montréal (CHUM). The collected blocks contain both invasive and in situ carcinoma of breast and normal breast tissue from healthy women undergoing plastic surgery.

Table S1D – China Cohort

A cohort of a total of 63 primary breast cancer and 22 benign breast tissue samples derived from 85 patients who underwent surgery at the First Affiliated Hospital of Anhui Medical University (Hefei, Anhui, People's Republic of China) between 2009 and 2010 was obtained. All tissue samples were hematoxylin and eosin stained and had been reviewed by two independent pathologists in Anhui Medical University. Total RNA from these breast tumor tissue samples was extracted by TRIzol® (Invitrogen, Life Technologies, Grand Island, NY), reversely transcribed into cDNA by using RevertAid First Strand cDNA Synthesis Kit (K1622, Fermentas, Germany) and Real-time PCR was

carried out by using SYBR Premix Ex Taq (DRR041A, TaKaRa, Shiga, Japan) in a Stratagene MX3000P detection system (Stratagene, La Jolla, CA, USA). The amplification protocol was set as following: an initial 95°C for 5min and then 40 cycles of denaturation at 95°C for 5sec, annealing and extension at 60°C for 30sec. Primer sequences used for qPCR expression in patient tissues from First Affiliated Hospital of Anhui Medical University are DP103-F: 5'-TGCCAGTAAACAGATGC-3', DP103-R: 5'-GTGCCAAAGGGTATGA-3'; MMP9-F: 5'-CGAACTTTGACAGCGACAAGA-3', MMP9-R: 5'- AGGGCGAGGACCATAGAGG-3'; GAPDH-F: 5'-TGCACCACCAACTGCTTAGC-3', GAPDH-R: 5'-GGCATGGACTGTGGTCATGAG-3'; HMGB1-F: 5' TTGTCGGGAGGAGCATAA 3', HMGB1-R: 5' GGGCGATACTCAGAGCAGAA 3'; H2AFZ-F: 5' CAAGACAAAGGCGGTTTC 3', H2AFZ-R: 5' GCATTTCCTGCCAGTTCA 3'.

Immunohistochemistry

DP103, MMP9, and phospho-p65 (S276) antibodies were used for immunohistochemical staining of the TMA sections. Briefly, the TMA sections were deparaffinized in Clearene and rehydrated through a graded series of ethanol. Endogenous peroxidase activity was quenched with 3% hydrogen peroxide for 30 min. Antigen retrieval was carried out through heating in 10mM citrate buffer (pH 6.0) at 90-100^{o}C for 20 min. The sections were blocked with goat serum for 1hr prior to overnight incubation at 4^{o}C with the primary antibody (1:50 dilution). The staining pattern was visualized using the Dako Envision-HRP kit with diaminobenzidine as the substrate. Sections were counterstained using Shandon's hematoxylin.

Immunohistochemical evaluations

The intensity of the staining in the epithelial compartment of ductal tissues was noted as absent (0), weak (1+), moderate (2+), or strong (3+). Low expression of DP103, phospho-p65 and MMP9 was defined as staining intensity of 0 or 1+, whereas high expression denotes staining intensities of 2+ or 3+. Evaluation of the stained TMA sections was carried out by two independent blinded observers, including a trained histopathologist.

SUPPLEMENTAL FIGURES

Figure S1

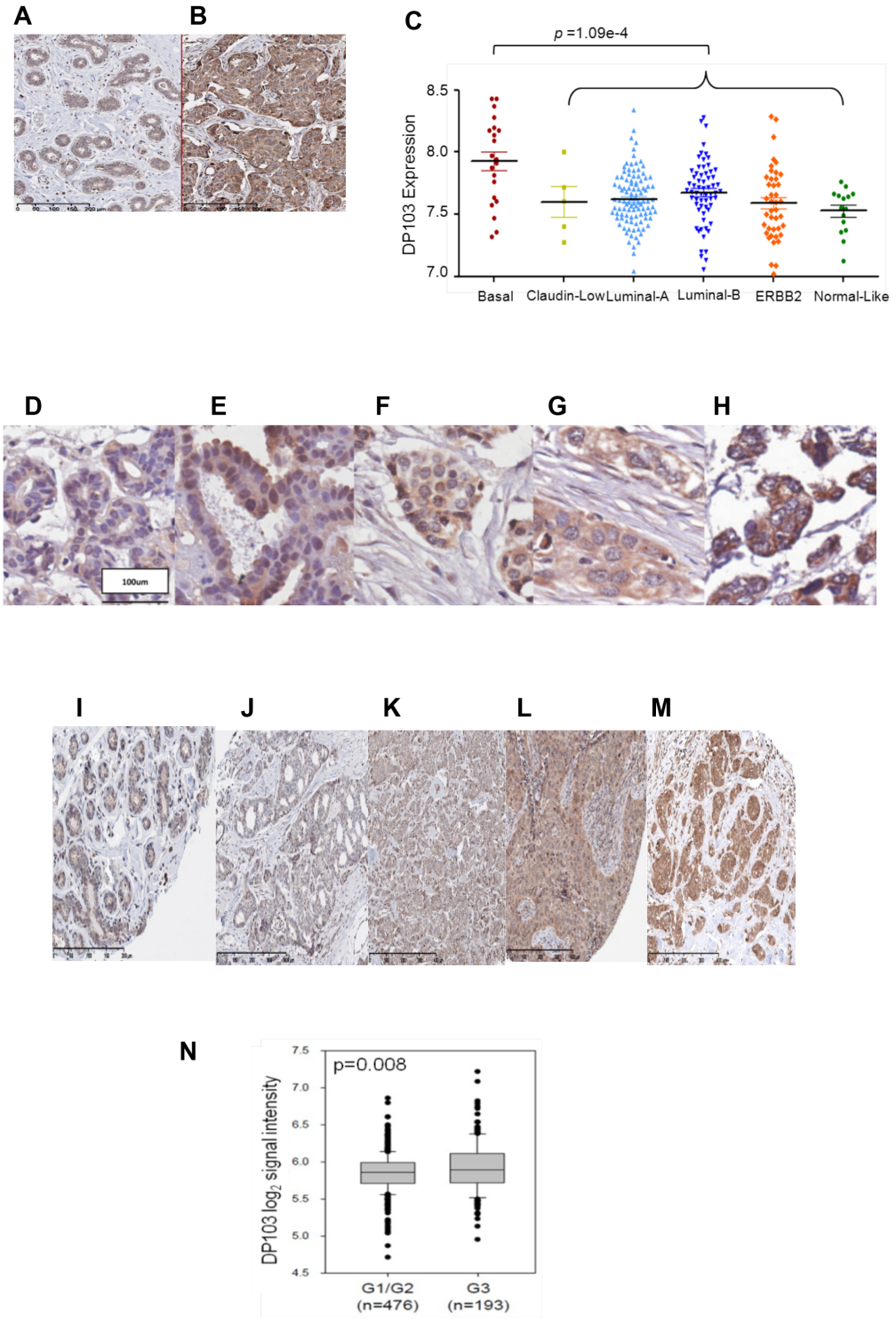
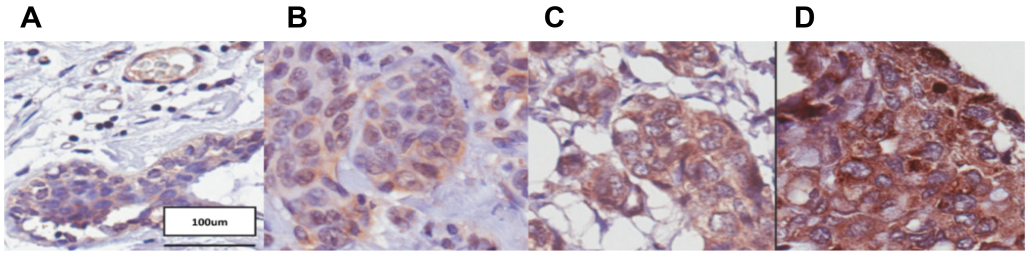
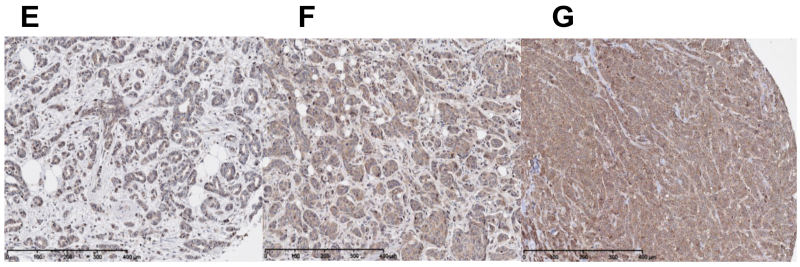


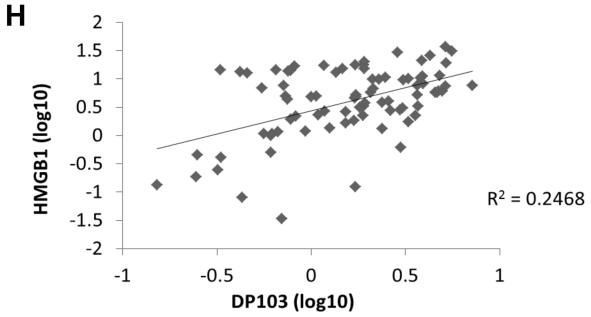
Figure S1 : Expression of DP103 in different molecular subtypes of breast cancer

DP103 staining in (A) a normal ductal tissue and (B) an invasive ductal carcinoma. (C) The gene expression value of DP103 (*y*-axis) is plotted for each breast cancer subtype. DP103 staining in (D) a normal ductal tissue, (E) in a Luminal A subtype, (F) in a Luminal B subtype, (G) in a HER2 subtype and (H) in a Basal subtype from cohort in Table S1A. DP103 staining in (I) a normal ductal tissue, (J) in a Luminal A subtype, (K) in a Luminal B subtype, (L) in a HER2 subtype and (M) in a Basal subtype from cohort in Table S1C. (N) Box and whisker plots of DP103 expression level distributions in tumors defined by histologic grade (Nottingham grading system). Shaded rectangles represent interquartile range; central line represents median value.

Figure S2


A
B
C
D
100um


E
F
G


H
HMGB1 (log10)
DP103 (log10)
R² = 0.2468
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-1
-0.5
0
0.5
1

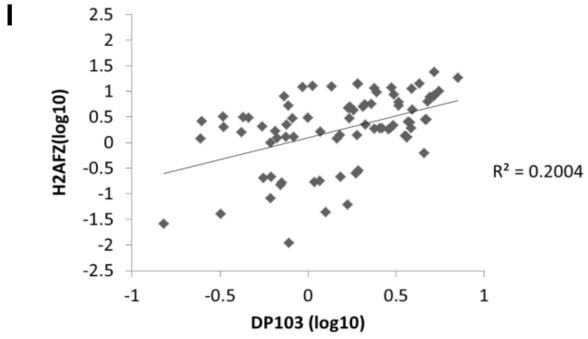

I
H2AFZ(log10)
DP103 (log10)
R² = 0.2004
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-1
-0.5
0
0.5
1

Figure S2: Expression of DP103 in different histological grades of breast cancer and its positive correlation with metastasis-associated genes.

DP103 staining (A) in normal non-malignant ductal tissue, (B) in low grade invasive ductal carcinoma, (C) in invasive ductal carcinoma of tumor grade 2 and 3 and (D) in high grade invasive ductal carcinoma from cohort in Table S1A. Cell nuclei were counterstained with haematoxylin. DP103 staining in (E) in grade I invasive ductal carcinoma, (F) in invasive ductal carcinoma of grade II tumor and (G) in grade III invasive ductal carcinoma from cohort in Table S1C. DP103 mRNA expression positively correlates to expression of metastasis genes (H) HMGB1 and (I) H2AFZ in 85 human breast patients from cohort in Table S1D showing positive correlation using Pearson's Correlation Coefficient.

Figure S3

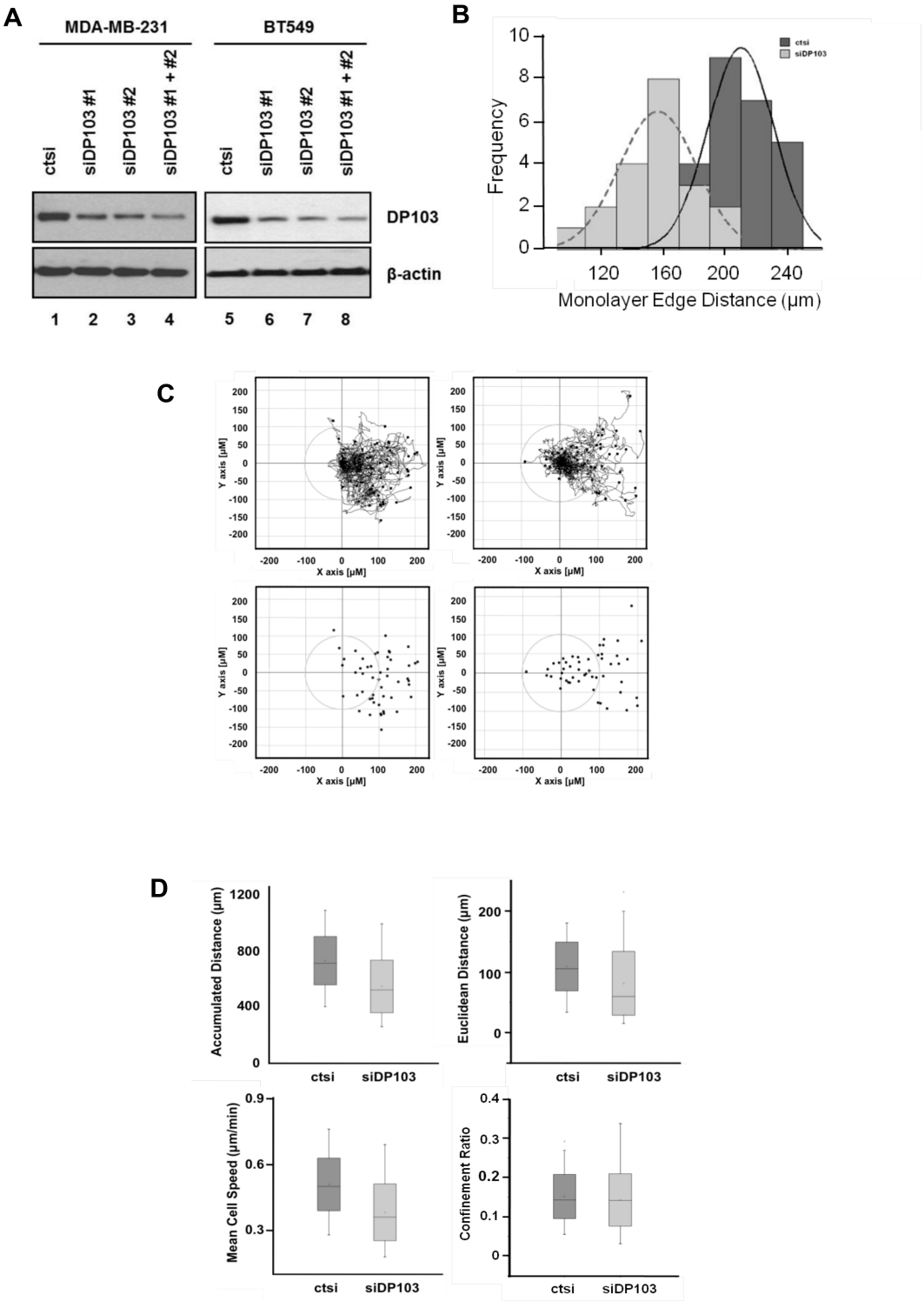

A
MDA-MB-231
BT549
ctsi
siDP103 #1
siDP103 #2
siDP103 #1 + #2
DP103
β-actin
1 2 3 4 5 6 7 8
B
Frequency
Monolayer Edge Distance (μm)
ctsi
siDP103
C
Y axis [μM]
X axis [μM]
D
Accumulated Distance (μm)
Euclidean Distance (μm)
Mean Cell Speed (μm/min)
Confinement Ratio
ctsi
siDP103

Figure S3: Suprresion of DP103 decreases migratory ability of breast cancer cells

(A) Transfection efficiencies of the two different siRNAs and in combination on expression of DP103 in MDA-MB-231 and BT549 cells. (B) MDA-MB-231 cells were subjected to siRNA knockdown of DP103 and the monolayer edge distances of the cells were tracked using live microscopy. (C) 2D plots of individual cell migration tracks with starting point of all the cell tracks coincided at the origin. Top and bottom left: control siRNA (ctsi) treated MDA-MB-231 show increased migratory distances with 72% of cells outside a radius of 101μm that coincides with the centre of mass (positive sign) of all the cell coordinates. Top and bottom right: siRNA against DP103 treated MDA-MB-231 show decreased migratory distances with 46% cell inside a radius of 74μm that coincides with the centre of mass of all the cell coordinates, and 56% of cells inside a radius of 101μm in comparison to ctsi. X and Y axes represent migration distance in μm. (D) Analyses of cell migration tracks. ctsi show a significant increase in the Accumulated distance (top left), Euclidean distance (top right), mean cell speed (bottom left) and confinement ratios (bottom right) in comparison to siDP103 cells. Box edges represent standard deviation; the small square within the box represents the mean and divider at 50% value.

Figure S4

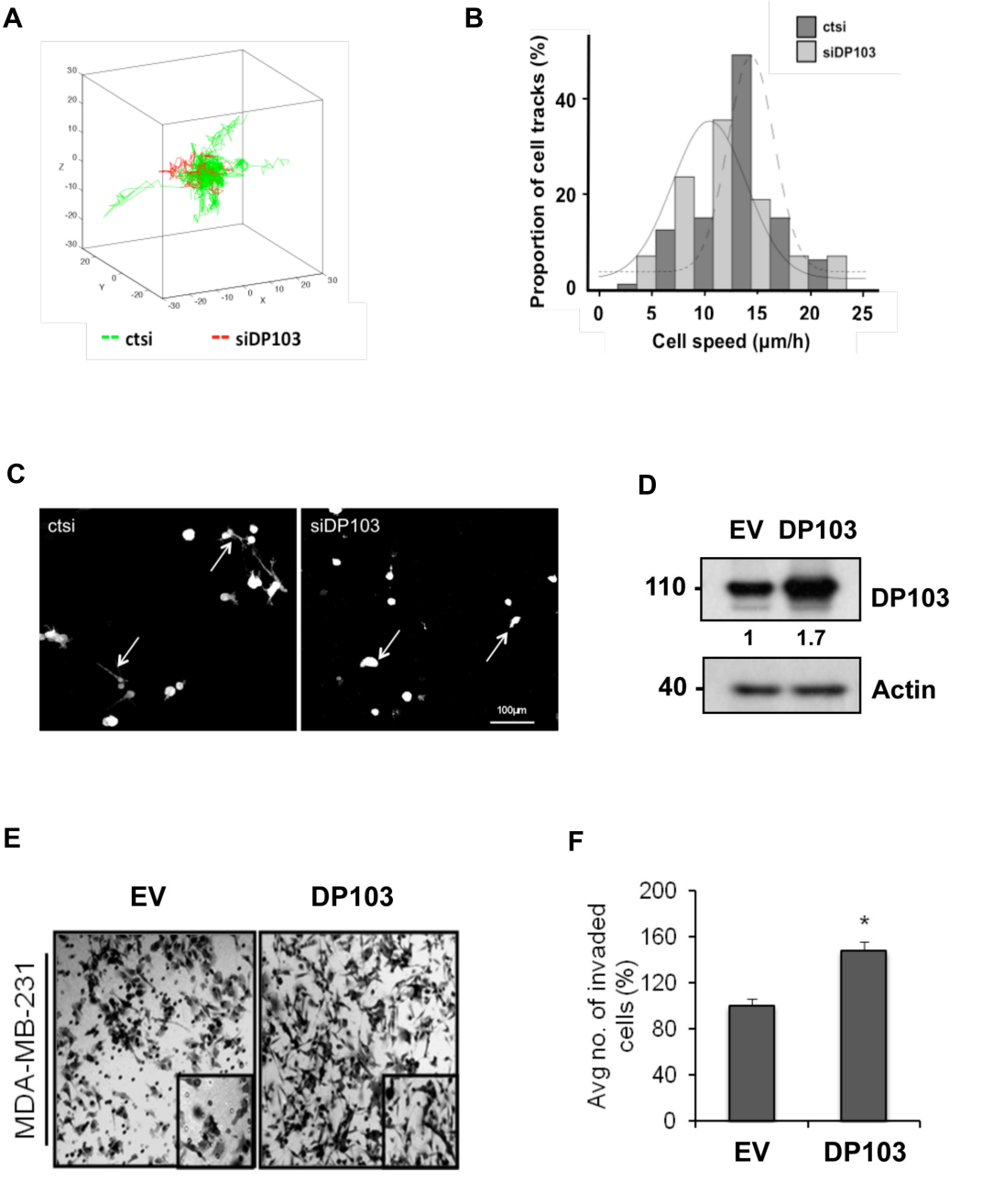

A
ctsi
siDP103
B
ctsi
siDP103
Proportion of cell tracks (%)
Cell speed (μm/h)
C
ctsi
siDP103
100μm
D
EV DP103
110
DP103
1 1.7
40
Actin
E
EV
DP103
MDA-MB-231
F
Avg no. of invaded cells (%)
EV
DP103

Figure S4: Suppression of DP103 decreases invasive ability of breast cancer cells while its ectopic expression enhances cell invasion

(A) Cell displacement (μm) tracks in 3D collagen gel. The tracks of a population of cells (n>50) were adjusted to start from the same origin (0,0,0). (B) The histogram of cells speed averaged over each track, track number >50. The height of each column corresponds to the percentage of tracks of a certain speed. (C) 3D Z-stack of confocal images was projected to X-Y plane, and cell morphology was shown with fluorescence from GFP. Left: Cells transfected with control siRNA (ctsi); Right: DP103 knockdown cells. Loss of pseudopodial protrusions in cells after siDP103 (white arrows). (D) MDA-MB-231 cells were transfected with pcDNA3/EV and pcDNA3-FLAG-DP103 (DP103/F-DP103). Cells extract immunoblotted with anti-DP103 antibodies. (E) MDA-MB-231 cells were transfected as stated in D and prepared for invasion assay. (F) MDA-MB-231 cells that invaded through the chambers counted and represented in percentages * denotes $p<0.05$.

Figure S5

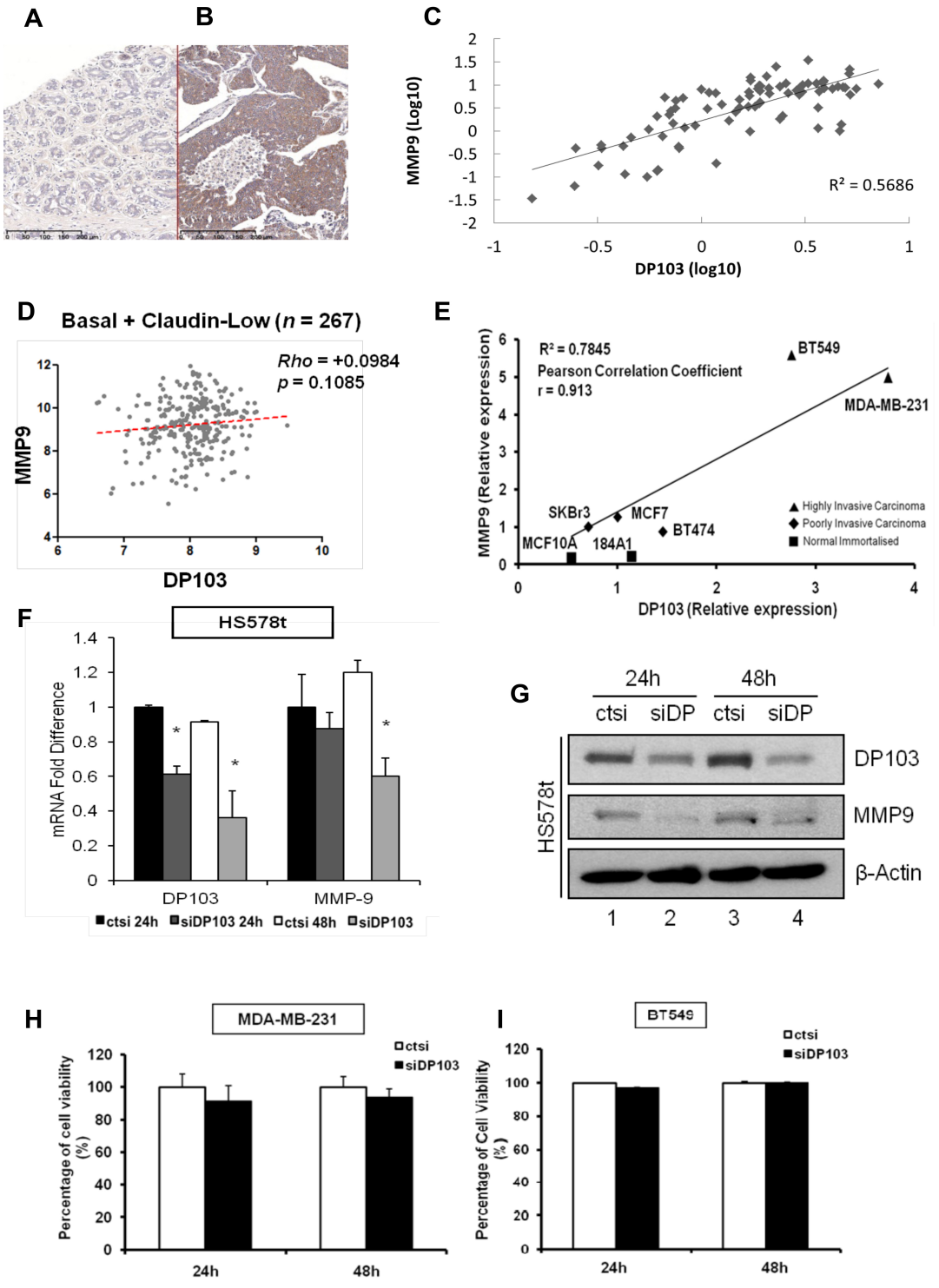

A
B
C
MMP9 (Log10)
DP103 (log10)
R² = 0.5686
D
Basal + Claudin-Low (n = 267)
Rho = +0.0984
p = 0.1085
MMP9
DP103
E
R² = 0.7845
Pearson Correlation Coefficient
r = 0.913
BT549
MDA-MB-231
SKBr3
MCF7
BT474
MCF10A
184A1
Highly Invasive Carcinoma
Poorly Invasive Carcinoma
Normal Immortalised
MMP9 (Relative expression)
DP103 (Relative expression)
F
HS578t
mRNA Fold Difference
DP103
MMP-9
ctsi 24h
siDP103 24h
ctsi 48h
siDP103
G
24h
48h
ctsi
siDP
ctsi
siDP
HS578t
DP103
MMP9
β-Actin
1
2
3
4
H
MDA-MB-231
ctsi
siDP103
Percentage of cell viability (%)
24h
48h
I
BT549
ctsi
siDP103
Percentage of Cell Viability (%)
24h
48h

Figure S5: MMP9 predict survival and strongly correlate with DP103 levels in breast cancer. Also, suppression of DP103 expression decreases MMP9 gene expression

(A-B) Different expression levels of MMP9 between (A) normal ductal tissue and (B) invasive ductal carcinoma tissue. (C) DP103 and MMP9 mRNA expression levels in 85 human breast patients were determined and showed positive correlation using Pearson's Correlation Coefficient (r^2=0.5686). (D) DP103 mRNA expression levels correlates positively with MMP9 mRNA expression in basal and claudin-low subtypes. (E) Pearson's Correlation Coefficient was determined between the mRNA expression of DP103 and MMP9 in various breast cell lines. DP103 knockdown in another invasive breast cancer cell line, HS578t, showed drop in (F) mRNA levels of MMP9 and (G) MMP9 protein levels * denotes $p<0.05$. (H) MDA-MB-231 and (I) BT549 cells transfected with either control siRNA or siDP103. Cell viability determined by crystal violet assay.

Figure S6

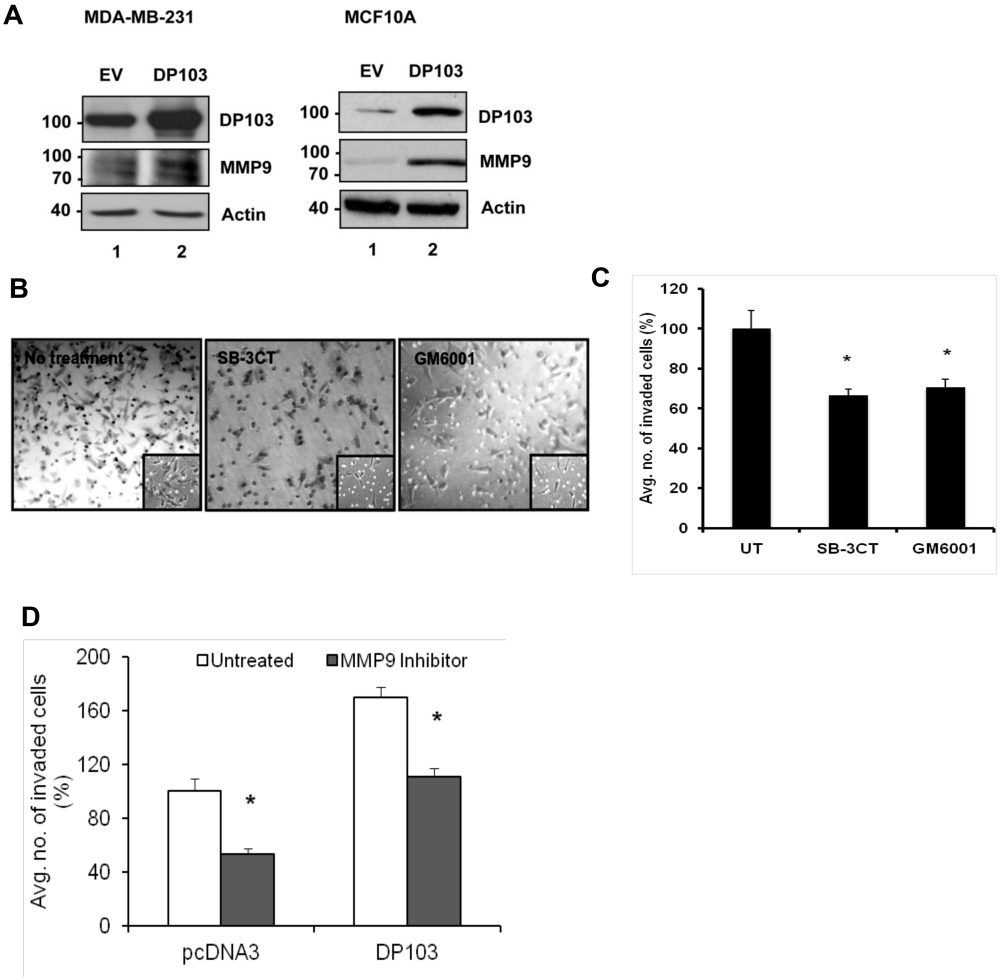


Figure S6: MMP9 mediates effects of DP103 in invasiveness

(A) MDA-MB-231 and MCF10A cells were transfected with pcDNA3 (EV) and pcDNA3-FLAG-DP103 (DP103). Cells extract immunoblotted with anti-DP103 and anti-MMP9 antibodies. (B) MDA-MB-231 cells assayed for invasion capability with either MMP2/9 inhibitor (SB-3CT, 10μM) or pan-MMPs inhibitor (GM6001, 50μM) (10X magnification). Insert shows zoomed pictures (40X magnification). (C) The number of cells invaded through the transwell invasion chambers from (B) counted and represented * denotes $p<0.05$. (D) MDA-MB-231 cells transfected with either empty vector (pcDNA3) or DP103 (DP103) assayed for transwell invasion with or without MMP9 Inhibitor I (1μM). The number of cells that invaded through the transwell invasion chambers counted and represented * denotes $p<0.05$.

Figure S7

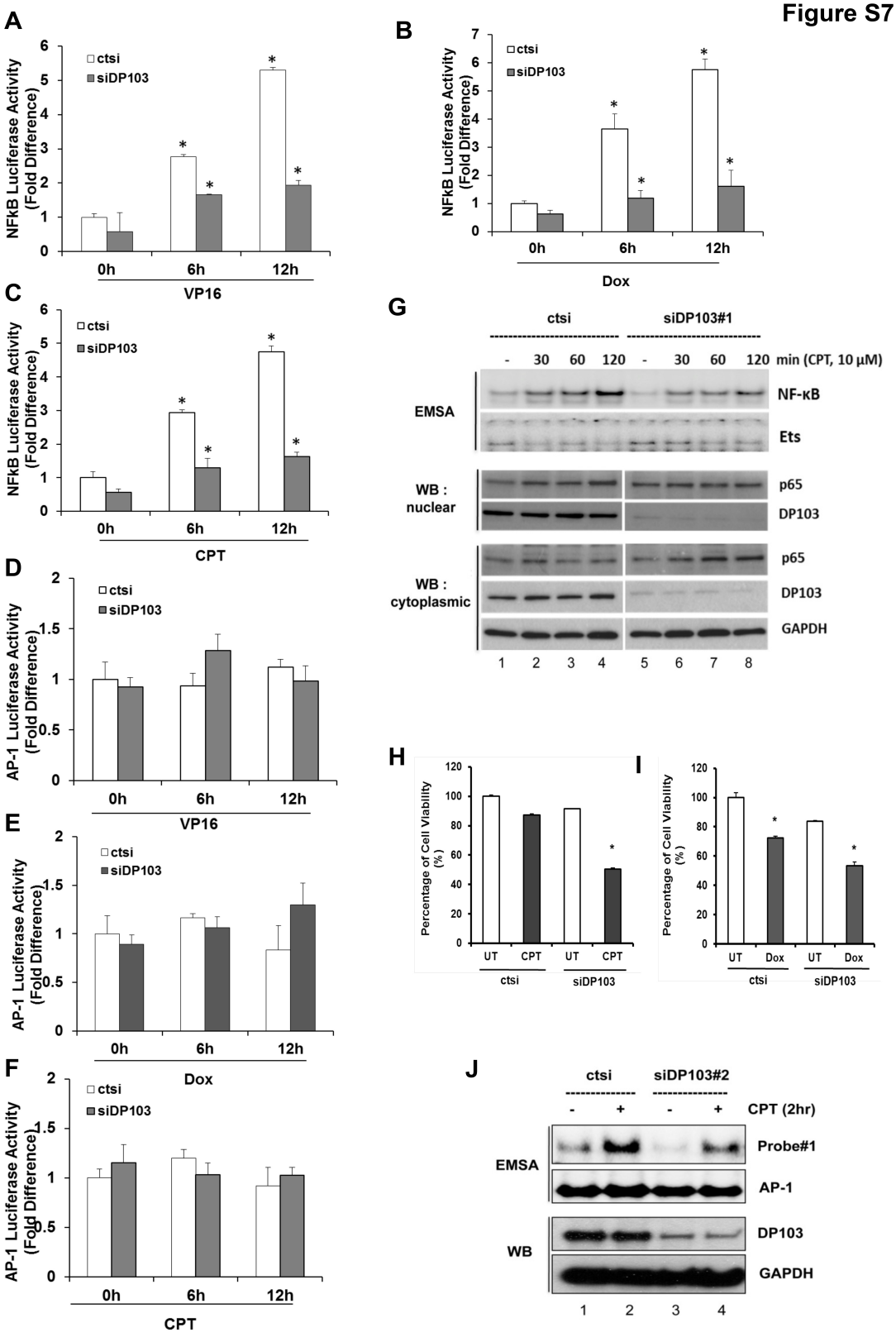

A
NFkB Luciferase Activity (Fold Difference)
ctsi
siDP103
0h
6h
12h
VP16
B
Dox
C
CPT
D
AP-1 Luciferase Activity (Fold Difference)
E
F
G
ctsi
siDP103#1
- 30 60 120 - 30 60 120 min (CPT, 10 μM)
EMSA
NF-κB
Ets
WB : nuclear
p65
DP103
WB : cytoplasmic
GAPDH
1 2 3 4 5 6 7 8
H
Percentage of Cell Viability (%)
UT
CPT
I
Dox
J
siDP103#2
CPT (2hr)
Probe#1
AP-1
WB
1 2 3 4

Figure S7: DP103 regulates NFκB-dependent gene expression

Control siRNA and siDP103 treated BT549 cells transfected with *Renilla* and luciferase reporter plasmid containing NFκB. The cells were subsequently stimulated with (A) 10μM VP16, (B) 25μM doxorubicin and (C) 10μM CPT for 0h, 6h, and 12h and then harvested for luciferase assay. Results are expressed in fold difference and are the average of three separate experiments * denotes $p<0.05$. Control siRNA and siDP103 treated BT549 cells were transfected with *Renilla* and luciferase reporter plasmid containing AP-1. The cells were subsequently stimulated with (D) 10μM VP16, (E) 25μM doxorubicin and (F) 10μM CPT for 0h, 6h, and 12h and then harvested for luciferase assay. Results are expressed in fold difference and are the average of three separate experiments * denotes $p<0.05$. (G) MDA-MB-231 cells transfected with control siRNA or siDP103. Cells were left untreated or treated with CPT (10μM) for indicated times. Protein extracts were examined with (top) EMSA and (bottom) Western blotting of nuclear and cytoplasmic protein fractions using antibody human-anti-p65, human-anti-DP103 and human-anti-GAPDH. (H) MDA-MB-231 cells were transfected with either control siRNA or DP103 siRNA and treated with or without (H) 10μM CPT or (I) 25μM doxorubicin for 48h. Graph showing percentage of cell viability (bottom) from three separate experiments * denotes $p<0.05$. (J) MDA-MB-231 cells transfected with control siRNA or siDP103. Cells were either left untreated or treated with CPT as indicated. Protein extracts were examined with EMSA and Western blotting using human-anti-DP103 and human-anti-GAPDH antibodies.

Figure S8

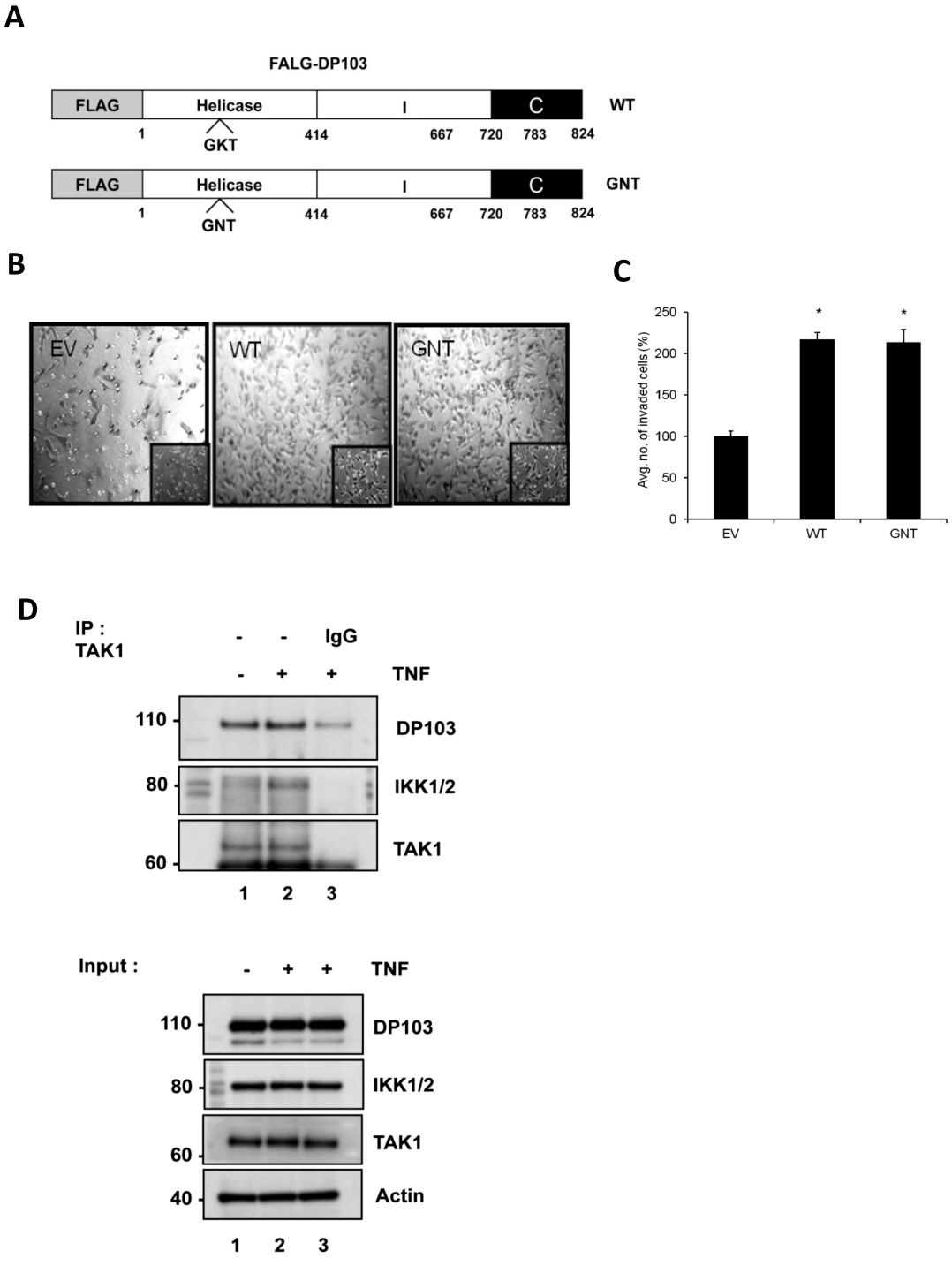

A
FALG-DP103
FLAG
Helicase
I
C
WT
1
GKT
414
667
720
783
824
FLAG
Helicase
I
C
GNT
1
GNT
414
667
720
783
824
B
EV
WT
GNT
C
Avg. no. of invaded cells (%)
250
200
150
100
50
0
*
*
EV
WT
GNT
D
IP :
TAK1
- - IgG
- + + TNF
110
DP103
80
IKK1/2
60
TAK1
1 2 3
Input :
- + + TNF
110
DP103
80
IKK1/2
60
TAK1
40
Actin
1 2 3

Figure S8: The helicase dead mutant GNT of DP103 retained its ability to induce invasion in MDA-MB-231 cells

(A) Map showing the FLAG-tagged wild-type (WT) and GNT mutant with single amino acid mutated in the helicase domain of DP103. (B) MDA-MB-231 cells transfected with either empty vector (EV), wild-type DP103 (WT) or GNT mutant (GNT) and assayed in transwell invasion chamber for 48h. Insert shows zoomed pictures. (C) The number of cells that invaded through the transwell invasion chambers in (B) counted and represented * denotes $p<0.05$. (D) MDA-MB-231 cells transfected with pGNT mutant were stimulated with TNFα for 10min. 0.5mg of protein lysate were immunoprecipitated with TAK1 and IgG antibodies and IP material and lysates were analysed by immunoblotting with the indicated antibodies. Western blot analysis of the input lysate is shown in the panel below.

Figure S9

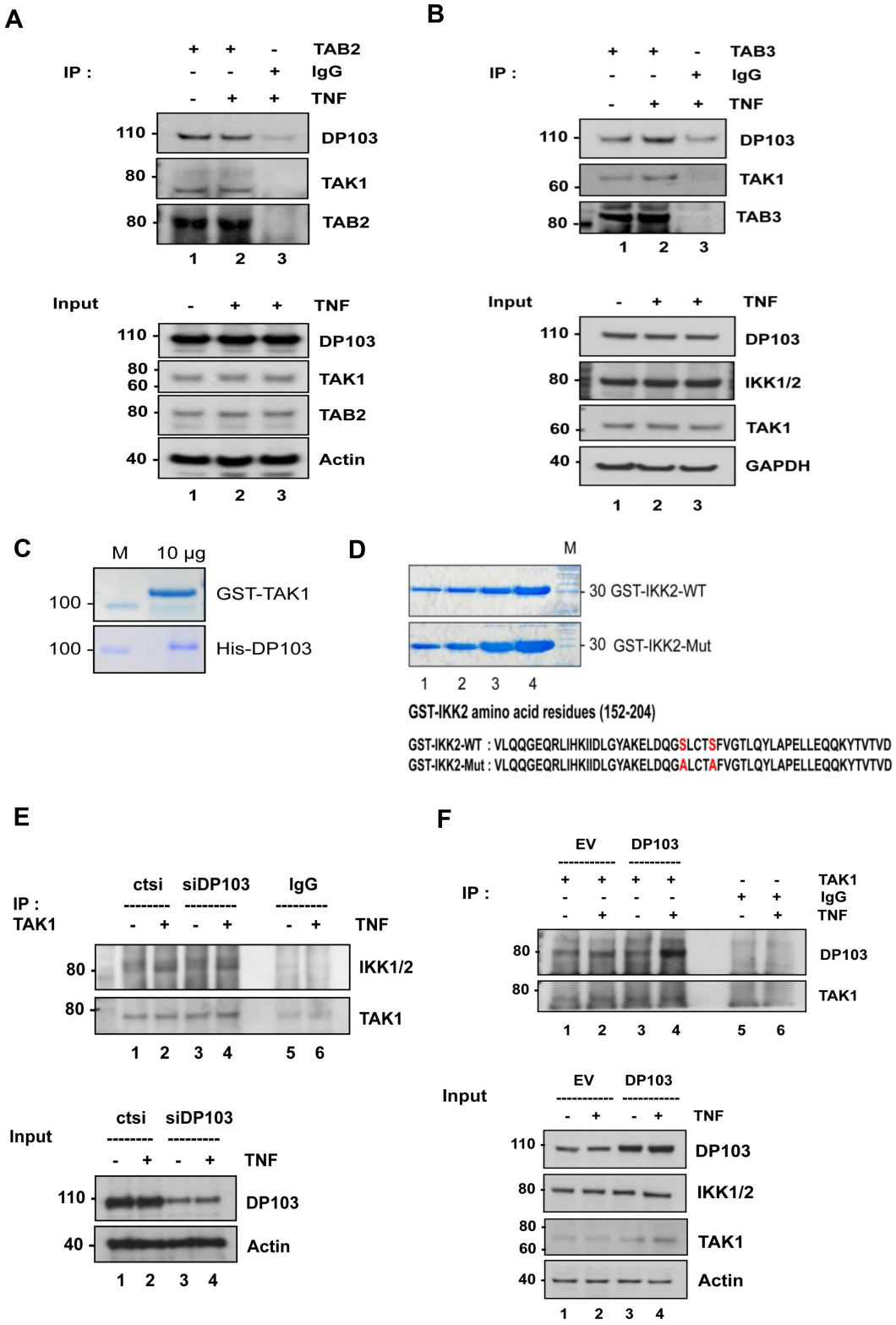

A
IP :
TAB2
IgG
TNF
110
DP103
80
TAK1
80
TAB2
1 2 3
Input
TNF
110
DP103
80
60
TAK1
80
TAB2
40
Actin
1 2 3
B
IP :
TAB3
IgG
TNF
110
DP103
60
TAK1
80
TAB3
1 2 3
Input
TNF
110
DP103
80
IKK1/2
60
TAK1
40
GAPDH
1 2 3
C
M
10 µg
100
GST-TAK1
100
His-DP103
D
M
30 GST-IKK2-WT
30 GST-IKK2-Mut
1 2 3 4
GST-IKK2 amino acid residues (152-204)
GST-IKK2-WT : VLQQGEQRLIHKIIDLGYAKELDQGSLCTSFVGTLQYLAPELLEQQKYTVTVD
GST-IKK2-Mut : VLQQGEQRLIHKIIDLGYAKELDQGALCTAFVGTLQYLAPELLEQQKYTVTVD
E
IP :
TAK1
ctsi
siDP103
IgG
TNF
80
IKK1/2
80
TAK1
1 2 3 4 5 6
Input
ctsi
siDP103
TNF
110
DP103
40
Actin
1 2 3 4
F
EV
DP103
IP :
TAK1
IgG
TNF
80
DP103
80
TAK1
1 2 3 4 5 6
Input
EV
DP103
TNF
110
DP103
80
IKK1/2
80
60
TAK1
40
Actin
1 2 3 4

Figure S9: DP103 enhances TAK1-mediated IKK2 phosphorylation and hence NFκB activation

(A) MDA-MB-231 cells were stimulated with TNFα for 10min. 0.5mg of protein lysate were immunoprecipitated with TAB2 and IgG antibodies and IP material and lysates were analysed by immunoblotting with the indicated antibodies. (B) MDA-MB-231 cells were stimulated with TNFα for 10min. 0.5mg of protein lysate were immunoprecipitated with TAB3 and IgG antibodies and IP material and lysates were analysed by immunoblotting with the indicated antibodies. (C) Purified recombinant full-length GST-TAK1 and His-DP103 proteins were stained using Coomassie blue staining. (D) Coomassie blue staining of purified GST-IKK2- WT and GST-IKK2-Mut proteins. The sequences are shown in the bottom panel and the amino acids in red indicate the wild type and mutant TAK1 phosphorylation sites in the activation loop of IKK2. (E) MDA-MB-231 cells transfected with siRNA control or siRNA against DP103 were stimulated with TNFα for 10min. 0.5mg of protein lysate were immunoprecipitated with TAK1 and IgG antibodies and IP material and lysates were analysed by immunoblotting with the indicated antibodies. Western blot analysis of the input lysate is shown in the panel below. (F) MDA-MB-231 cells infected with empty vector (EV) or pBOBI-DP103 were stimulated with TNFα for 10min. IP material and lysates were analysed by immunoblotting with the indicated antibodies as described in (E).

SUPPLEMENTAL TABLES

Table S1: A) & B) Clinicopathological features (Singapore)

A)

Clinicopathological features	Number of cases
Age (years)	
Mean	53
Median	51
Minimum	19
Maximum	86
Ethnicity	
Chinese	336
Malay	35
Indian	17
Others	11

NA= not available

B)

Clinicopathological features	Number of cases
Histological tumor grade	
1	66
2	151
3	170
NA	12
Tumor type	
Luminal A	175
Luminal B	38
Basal	60
HER2	46
NA	80

Table S1: C) Clinicopathological features (Canada)

Clinicopathological features	Number of cases
Age (35-92 years)	411
Grade	
I	38
II	70
III	303
Histological subtypes	
Invasive ductal carcinoma	326
Invasive lobular carcinoma	10
Typical medullary carcinoma	49
Atypical medullary carcinoma	19
Colloid carcinoma	7
Estrogen receptor	
Positive	94
Negative	225
Progestron receptor	
Positive	94
Negative	225
HER2 receptor	
Positive	63
Negative	256

Table S1: D) Clinicopathological features (China)

Clinicopathological features	Number of cases
Age (years)	
= 35	5
35-55	42
> 55	16
Tumor size (=cm)	
= 2	26
2-5	34
> 5	3
Lymph node metastasis	
Absent	33
Present	30
Grade	
I	14
II	37
III	12
Histological subtypes	
Ductal	60
Lobular	1
Others	2
Estrogen receptor	
-	19
+	44
Progestron receptor	
-	22
+	41
c-erbB-2 receptor	

Low	30
High	33
Ki-67	
+	43
++	16
+++	2
++++	2

Table S1. (A) Clinicopathological features of number of cases from Singapore cohort. Collated expression analysis of DP103 epitope in the epithelial compartments of indicated number of ductal specimens from (B) Singapore cohort and (C) Canada cohort. (D) Clinicopathological features of number of cases from China cohort.

Table S2: A

	DP103 expression		
	Low	High	*P* value
Diagnosis			
Normal	38	0	< 0.001
Malignant	63	267	

Table S2: B

	DP103 expression		
	Low	High	*P* value
Diagnosis			
Basal	3	47	0.0053
Other tumor subtypes	50	164	

Table S2: C

	DP103 expression		
	Low	High	*P* value
Diagnosis			
Normal	38	0	< 0.001
1	51	5	
2	5	128	
3	0	133	

Table S2: D

	DP103 expression		
	Low	High	*P* value
Diagnosis			
Normal	15	0	< 0.001
Malignant	0	385	

Table S2: E

	DP103 expression		
	Low	High	*P* value
Triple negative	8	144	0.0001
Other tumor subtypes	24	157	

Table S2: F

	DP103 expression		
	Low	High	*P* value
Normal	15	0	
Grade I	2	24	< 0.001
Grade II	0	47	
Grade III	27	242	

Table S2. Collated expression analysis of DP103 epitope in basal and non-basal breast tumor subtypes in specimens from (A) Singapore cohort and (B) Canada cohort. Association analyses on expression of DP103 in the various tumor grades were determined using Fisher's Exact and Kendall-Tau tests in specimens from (C) Singapore cohort and (D) Canada cohort. Analysis of the tissue microarrays showed that the staining intensity of DP103 in the epithelial compartment was highly significantly associated with histological tumor grade.

Table S3:

Cell line	No. of mouse with metastasis	Metastasis (p/s)	
		Lung	Liver
MDA-MB-231-DP103	4/11	7.9×10^7 1.0×10^6 1.3×10^6 3.8×10^6	5.4×10^7 8.2×10^7 7.0×10^6 1.3×10^8
MDA-MB-231-EV	1/12	1.9×10^7	2.3×10^7

Table S3. Spontaneous metastases from orthotopic mammary fat pad implant. Two million cells empty vector transfected MDA-MB-231 cells (MDA-MB-231-EV) or DP103 transfected cells (MDA- MB-231-DP103) are injected orthotopically into the abdominal m.f.p. of female nude mice. The incidence of pulmonary and liver metastases at necropsy from primary tumors formed by these two cancer cell lines was evaluated by bioluminescence imaging as described in methods and tabulated.

Table S4: A

	MMP9 expression		
	Low	High	*P* value
Diagnosis			
Normal	53	0	< 0.001
Malignant	71	286	

Table S4: B

	MMP9 expression		
	Low	High	*P* value
Diagnosis			
Normal	15	0	< 0.001
Malignant	46	339	

Table S4: C

	DP103 expression		
	Low	High	*P* value
MMP9 expression			
Low	56	10	< 0.001
High	6	253	

Table S4: D

	DP103 expression		
	Low	High	*P* value
MMP9 expression			
Low	4	74	< 0.001
High	19	196	

Table S4. Collated expression pattern analysis of MMP9 epitope in epithelial compartments of ductal specimens from (A) Singapore cohort and (B) Canada cohort. Normal non-malignant ductal tissues had reduced staining in comparison to the malignant invasive ductal carcinoma tissues. Association analysis between expression of DP103 and MMP9 were determined using Fisher's Exact and Kendall-Tau tests in specimens from (C) Singapore cohort and (D) Canada cohort. Results suggested that high DP103 expression is highly significantly correlated with high MMP9 levels.

Table S5: A

	Phospho-p65 expression		
	Low	High	*P* value
Diagnosis			
Normal	49	0	< 0.001
Malignant	69	269	

Table S5: B

	Phospho-p65 expression		
	Low	High	*P* value
Diagnosis			
Basal	5	47	0.0141
Other tumor subtypes	50	141	

Table S5: C

	Phospho-p65 expression		
	Low	High	*P* value
Histological tumor grades			
Normal	49	0	< 0.001
1	54	4	
2	4	130	
3	4	134	

Table S5: D

	DP103 expression		
	Low	High	*P* value
Phospho-p65 expression			
Low	55	10	< 0.001
High	7	246	

Table S5. (A) Collated expression analysis of phospho-p65 S276 epitope in the epithelial compartments of ductal specimens. Normal non-malignant ductal tissues had reduced staining in comparison to the malignant invasive ductal carcinoma tissues. (B) Collated expression analysis of phospho-p65 S276 epitope and its association with tumor subtype. High expression of phospho-p65 S276 was significantly associated with basal breast tumor subtype. (C) Collated expression analysis of phospho-p65 S276 epitope and its association with histological tumor grades. Analysis of the tissue microarrays showed that the staining intensity of phospho-p65 in the epithelial compartment was highly significantly associated with histological tumor grade. (D) Association analysis between expression of DP103 and phospho-p65 (Ser276) were determined using Fisher's Exact and Kendall-Tau tests. Results suggest that high DP103 expression is highly significantly correlated with high phospho-p65 levels

Appendix IV: Identification of genes that could distinguish between luminal A and luminal B molecular subtypes of breast cancer using *in silico* analysis

Our objective was to find out new biomarkers that could facilitate our ability to distinguish between luminal A and luminal B subtypes of breast cancer. *In silico* analysis was carried out using RNA sequencing data of 756 patients to identify genes whose expression pattern varied differentially between these two subsets of tumors. As shown in Figure 1, there were 20532 genes that significantly (< 0.05) have different expression between luminal A and luminal B breast cancer (with higher expression in luminal B tumors). Only, 1966 genes could reach to the < 0.01 significant level, with higher expression in luminal B tumors. From these genes, there are 162 genes whose expression has no overlap between boxes in the boxplots (Table 1). However, the number of those genes is not applicable to be studied using IHC and TMA of human tissue. We tried to apply certain criteria to select the best candidates for our study that includes; selection of proteins that are highly expressed in luminal B than in luminal A based on the web application bc-GenExMiner [415], positive correlation with Ki-67, inclusion in the proliferation peak according to MiSTIC, established role in proliferation and localization only to the nuclei (Literature, Human Protein Atlas and genecard website). Based on this short list, only 19 genes fulfilled the abovementionned criteria (Table 2) and are supposedly putative targets to readily distinguish between luminal A and luminal B breast cancers (Figure 2).

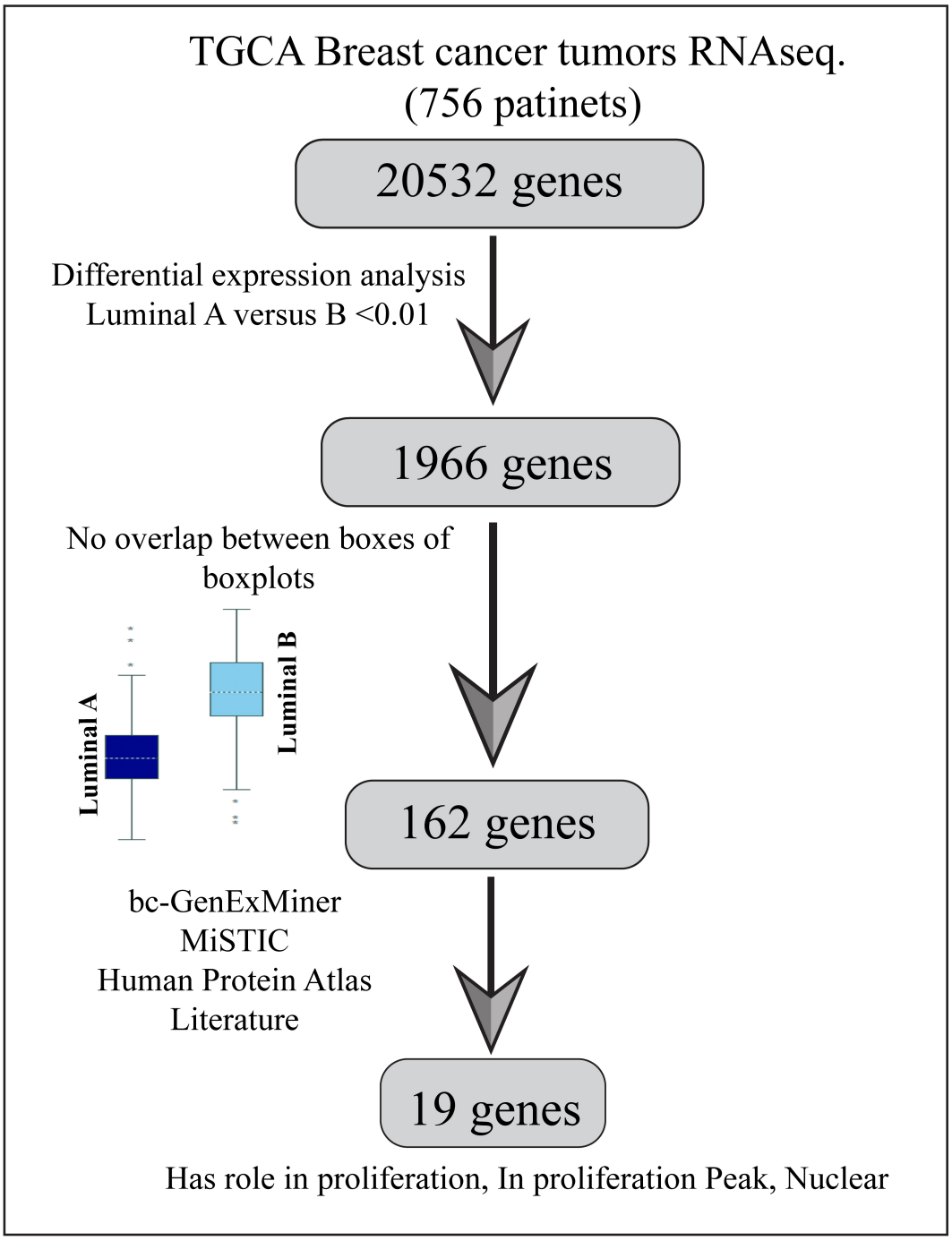
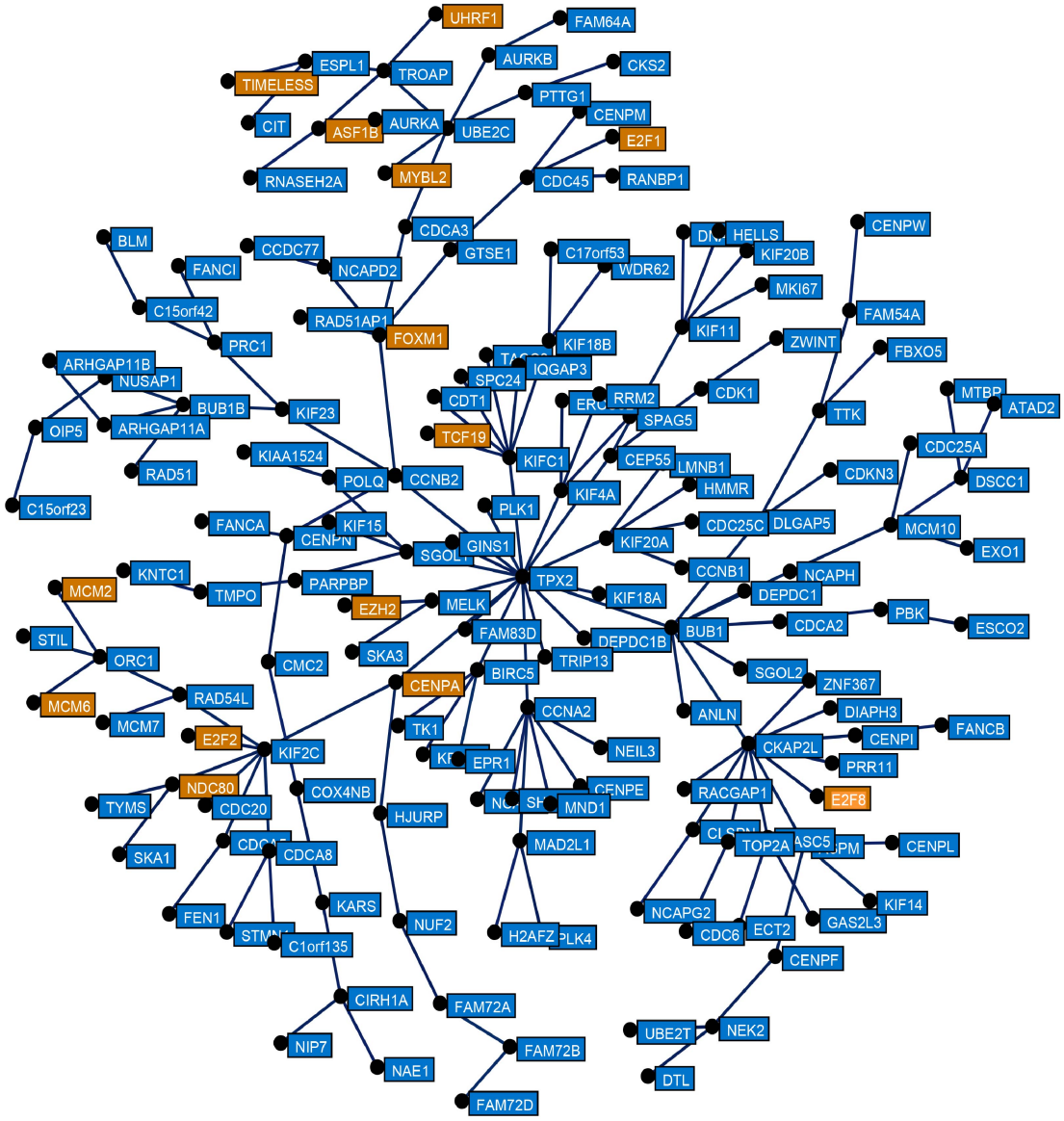


Figure 1: steps of selection of target genes that can distinguish between luminal A and luminal B breast cancer



CopySave as

Search: proliferation

P-value	Odds	Name	Type	Cat	ID
3.80e-156	438.78	ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER	msigdb	C2/CGP	M1
1.70e-136	422.67	module #11 Prolif: cell proliferation	microarrays	Breast Cancer Co-Expression module	46
8.90e-135	75.94	proliferation signature (Venet D et al. 2011)	microarrays	Signature	389
6.74e-94	120.98	CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP	msigdb	C2/CGP	M3
2.56e-76	133.71	genes most positively correlated with proliferation marker PCNA (Venet D et al. 2011)	microarrays	Signature	388
2.45e-71	468.24	Breast Cancer molecular subtype [CIT] - cluster b (cycle/proliferation)	microarrays	Signature	64
4.22e-71	97.75	BENPORATH_PROLIFERATION	msigdb	C2/CGP	M2
6.47e-14	41.19	Breast Cancer Intrinsic + Proliferation gene set - Perreard et al., 2006	microarrays	Signature	249
1.82e-13	10.37	cell proliferation	go	BP	000
3.80e-10	26.33	PEART_HDAC_PROLIFERATION_CLUSTER_UP	msigdb	C2/CGP	M1

Showing 1 to 10 of 17 entries (filtered from 3,557 total entries)

FirstPreviousNextLast

Figure 2: Proliferation cluster according to MiSTIC that include some of the genes of interest.

Table 1: List of genes that are significantly highly expressed in luminal B than in luminal A breast cancer

No.	Gene	Description	Gene ID	BaseMean LumB	BaseMean LumA	*P value*
1	MYBL2	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	4605	22.5612802	5.45347035	2.03E-47
2	UBE2C	Ubiquitin-conjugating enzyme E2C	11065	54.0548116	16.3211930	1.15E-61
3	BIRC5	Baculoviral IAP repeat containing 5	332	57.4202590	17.7166410	1.56E-40
4	TOP2A	Topoisomerase (DNA) II alpha 170kDa	7153	83.4312383	25.9139120	1.93E-28
5	CDC6	Cell division cycle 6 homolog (S. cerevisiae)	990	12.2336271	4.15630539	4.20E-12
6	TROAP	Trophinin associated protein (tastin)	10024	9.51400063	3.26005655	1.73E-36
7	AURKA	Aurora kinase A	6790	15.3307455	5.41037698	2.48E-47
8	TPX2	TPX2, microtubule-associated, homolog (Xenopus laevis)	22974	25.2692678	8.93857292	2.32E-62
9	CCNB2	Cyclin B2	9133	21.0276064	7.57380166	4.67E-56
10	FOXM1	Forkhead box M1	2305	15.8692775	5.73305857	1.74E-41
11	CCNE2	Cyclin E2	9134	12.2881905	4.47509131	1.86E-33
12	EPR1	Effector cell peptidase receptor 1 (non-protein coding)	8475	5.59553581	2.03984849	3.24E-23
13	SPAG5	Sperm associated antigen 5	10615	13.4969697	4.97266578	1.19E-39
14	UBE2T	Ubiquitin-conjugating enzyme E2T (putative)	29089	30.7293649	11.4604247	4.94E-60
15	KIF2C	Kinesin family member	11004	10.5085973	3.92776939	3.71E-36

		2C				
16	CCNA2	Cyclin A2	890	11.1416815	4.16783656	3.85E-36
17	RRM2	Ribonucleotide reductase M2	6241	21.1899761	7.95771322	1.12E-39
18	AURKB	Aurora kinase B	9212	9.42374092	3.54873814	6.47E-34
19	PLK1	Polo-like kinase 1	5347	14.07631319	5.301641822	5.96E-42
20	MELK	Maternal embryonic leucine zipper kinase	9833	10.70792702	4.06248384	2.17E-35
21	CEP55	Centrosomal protein 55kDa	55165	11.13325208	4.231002183	9.11E-36
22	CDK1	Cyclin-dependent kinase 1	983	31.36632574	12.14120845	1.10E-40
23	PRC1	Protein regulator of cytokinesis 1	9055	25.33267512	9.888556646	1.13E-50
24	CDKN3	Cyclin-dependent kinase inhibitor 3	1033	19.73354066	7.71829765	1.99E-45
25	MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)	4085	20.79272789	8.153700774	1.33E-45
26	PBK	PDZ binding kinase	55872	10.44410264	4.096301906	6.77E-30
27	E2F8	E2F transcription factor 8	79733	3.514804923	1.380413666	1.91E-13
28	SGOL1	Shugoshin-like 1 (S. pombe)	151648	4.763050728	1.877912897	8.75E-18
29	CDC20	Cell division cycle 20 homolog (S. cerevisiae)	991	18.71899424	7.393666306	4.02E-41
30	CDC45	Cell division cycle 45 homolog (S. cerevisiae)	8318	5.953133949	2.351757406	4.41E-22
31	ASPM	asp (abnormal spindle) homolog, microcephaly associated (Drosophila)	259266	7.322075181	2.909558614	2.70E-24
32	MCM10	Minichromosome maintenance complex component 10	55388	3.760270946	1.499660267	1.14E-15
33	E2F1	E2F transcription factor 1	1869	8.971866952	3.588748062	3.61E-27

34	PTTG1	Pituitary tumor-transforming 1	9232	31.36775093	12.55534751	7.78E-46
35	CENPA	Centromere protein A	1058	5.887119358	2.360024949	2.95E-22
36	NUSAP1	Nucleolar and spindle associated protein 1	51203	26.46628428	10.61239344	1.05E-48
37	NDC80	NDC80 kinetochore complex component homolog (S. cerevisiae)	10403	8.536390669	3.426616685	1.51E-27
38	BUB1	Budding uninhibited by benzimidazoles 1 homolog (yeast)	699	9.605833444	3.861973889	2.92E-29
39	NCAPG	Non-SMC condensin I complex, subunit G	64151	6.685396906	2.68791162	5.18E-22
40	UHRF1	Ubiquitin-like with PHD and ring finger domains 1	29128	14.37347409	5.789620827	5.23E-35
41	MKI67	Antigen identified by monoclonal antibody Ki-67	4288	13.03307084	5.254331813	9.23E-35
42	KIF4A	Kinesin family member 4A	24137	8.296237485	3.345751033	1.72E-25
43	DLGAP5	Discs, large (Drosophila) homolog-associated protein 5	9787	7.691272319	3.122815093	2.11E-24
44	CKAP2L	Cytoskeleton associated protein 2-like	150468	5.00579807	2.032800472	4.13E-17
45	ANLN	Anillin, actin binding protein	54443	11.2754635	4.615518027	1.72E-32
46	BUB1B	Budding uninhibited by benzimidazoles 1 homolog beta (yeast)	701	10.62362119	4.356970525	1.13E-28
47	NEK2	NIMA (never in mitosis gene a)-related kinase 2	4751	18.36402133	7.568874618	5.94E-38
48	DSCC1	Defective in sister	79075	6.546108431	2.699219059	1.22E-20

		chromatid cohesion 1 homolog (S. cerevisiae)				
49	ESPL1	Extra spindle pole bodies homolog 1 (S. cerevisiae)	9700	3.928328133	1.622730444	3.33E-13
50	MND1	Meiotic nuclear divisions 1 homolog (S. cerevisiae)	84057	5.134850579	2.122949804	8.13E-18
51	TRIP13	Thyroid hormone receptor interactor 13	9319	12.26161151	5.102962031	1.13E-31
52	CDCA5	Cell division cycle associated 5	113130	11.59976315	4.848443452	6.46E-30
53	CDCA3	Cell division cycle associated 3	83461	7.015025423	2.932708801	1.41E-22
54	EXO1	Exonuclease 1	9156	5.503272414	2.302433579	1.29E-18
55	MLF1IP	MLF1 interacting protein	79682	16.50960816	6.931655656	2.37E-34
56	TTK	TTK protein kinase	7272	6.368888467	2.680333217	4.16E-21
57	NUF2	NUF2, NDC80 kinetochore complex component, homolog (S. cerevisiae)	83540	11.39017715	4.805649425	3.37E-29
58	CCNB1	Cyclin B1	891	32.63932898	13.77564215	4.79E-48
59	ATAD2	ATPase family, AAA domain containing 2	29028	23.52900372	9.933064828	1.13E-40
60	FAM83D	Family with sequence similarity 83, member D	81610	7.283453576	3.08062417	5.26E-22
61	CENPE	Centromere protein E, 312kDa	1062	4.066544804	1.727809448	2.36E-13
62	KPNA2	Karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	3838	95.81742474	40.77903671	3.47E-43
63	TYMS	Thymidylate synthetase	7298	26.76991293	11.39933121	1.99E-32
64	CDCA8	Cell division cycle associated 8	55143	10.20273559	4.352052726	4.89E-28
65	NCAPH	Non-SMC condensin I	23397	5.070918115	2.183032477	4.36E-16

		complex, subunit H				
66	DEPDC1	DEP domain containing 1	55635	4.153440889	1.80544395	2.44E-14
67	KIF23	Kinesin family member 23	9493	11.06854727	4.8225495	2.46E-26
68	ZWINT	ZW10 interactor	11130	31.77903524	13.90331721	7.47E-44
69	KIF15	Kinesin family member 15	56992	4.735634414	2.073602085	1.56E-14
70	KIF20A	Kinesin family member 20A	10112	10.06915908	4.432234997	4.19E-25
71	KIF18A	Kinesin family member 18A	81930	5.926833373	2.611963871	7.56E-17
72	IQGAP3	IQ motif containing GTPase activating protein 3	128239	10.84228695	4.78452376	8.16E-26
73	RACGAP 1	Rac GTPase activating protein 1	29127	20.52655907	9.07106785	1.01E-34
74	DEPDC1 B	DEP domain containing 1B	55789	3.77042553	1.667463791	4.42E-12
75	HJURP	Holliday junction recognition protein	55355	6.222561501	2.75454221	2.81E-18
76	KIF11	Kinesin family member 11	3832	15.15731627	6.714665773	4.45E-30
77	CENPF	Centromere protein F, 350/400kDa (mitosin)	1063	13.90916199	6.163152483	1.16E-29
78	LMNB1	Lamin B1	4001	30.4946221	13.51946745	1.44E-42
79	KIFC1	Kinesin family member C1	3833	12.47527417	5.5340709	1.12E-28
80	GTSE1	G-2 and S-phase expressed 1	51512	9.986669508	4.43574954	4.95E-25
81	KIF18B	Kinesin family member 18B	146909	5.538756658	2.463040247	5.24E-16
82	GINS1	GINS complex subunit 1 (Psf1 homolog)	9837	7.609546683	3.386084893	1.42E-19

83	ARHGAP 11A	Rho GTPase activating protein 11A	9824	7.23887503	3.233067779	1.82E-18
84	FAM72D	Family with sequence similarity 72, member D	728833	3.344285675	1.497213937	3.84E-11
85	FAM72B	Family with sequence similarity 72, member B	653820	3.167482734	1.421347072	9.49E-11
86	RAD51	RAD51 homolog (S. cerevisiae)	5888	5.919803196	2.673817983	6.22E-16
87	RECQL4	RecQ protein-like 4	9401	7.538265069	3.406595705	1.55E-18
88	DIAPH3	Diaphanous homolog 3 (Drosophila)	81624	9.520660474	4.308330623	3.72E-20
89	FAM64A	Family with sequence similarity 64, member A	54478	4.792426209	2.183308365	1.38E-14
90	DTL	Denticleless E3 ubiquitin protein ligase homolog (Drosophila)	51514	10.27333209	4.681351781	1.50E-21
91	CDC25C	Cell division cycle 25 homolog C (S. pombe)	995	4.415865467	2.016897262	5.59E-12
92	HMMR	Hyaluronan-mediated motility receptor (RHAMM)	3161	6.941401232	3.176887115	4.00E-17
93	SHCBP1	SHC SH2-domain binding protein 1	79801	5.06372862	2.321716222	4.52E-14
94	FAM54A	Family with sequence similarity 54, member A	113115	3.098604	1.424061641	2.69E-10
95	MYBL1	v-myb myeloblastosis viral oncogene homolog (avian)-like 1	4603	5.505701383	2.530823817	2.40E-11
96	FAM72A	Family with sequence similarity 72, member A	729533	2.77178156	1.277734735	4.03E-09
97	RAD51A P1	RAD51 associated protein 1	10635	6.800423623	3.138928644	1.12E-17
98	RAD54L	RAD54-like (S.	8438	3.923622577	1.821366416	6.29E-12

		cerevisiae)				
99	KIF14	Kinesin family member 14	9928	2.818038653	1.309073217	6.60E-09
100	NEIL3	Nei endonuclease VIII-like 3 (E. coli)	55247	2.920613807	1.358010886	6.39E-09
101	KIAA0101	KIAA0101	9768	17.02374109	7.930428251	1.25E-27
102	GINS2	GINS complex subunit 2 (Psf2 homolog)	51659	25.27644279	11.85970134	2.36E-30
103	C15orf42	Chromosome 15 open reading frame 42	90381	2.434535576	1.144459147	4.71E-08
104	HMGB2	High mobility group box 2	3148	104.6331187	49.36988843	8.74E-43
105	SKA3	Spindle and kinetochore associated complex subunit 3	221150	4.456471849	2.116230181	2.68E-12
106	ASF1B	ASF1 anti-silencing function 1 homolog B (S. cerevisiae)	55723	17.22805226	8.252568134	2.16E-27
107	TACC3	Transforming, acidic coiled-coil containing protein 3	10460	20.09643355	9.630079451	3.96E-30
108	FANCA	Fanconi anemia, complementation group A	2175	8.235418606	3.947952427	5.87E-19
109	SPC24	SPC24, NDC80 kinetochore complex component, homolog (S. cerevisiae)	147841	4.032725109	1.936938501	1.10E-10
110	CENPK	Centromere protein K	64105	4.754524154	2.305517872	6.63E-11
111	FANCI	Fanconi anemia, complementation group I	55215	16.4236062	7.988116512	7.94E-25
112	RAD54B	RAD54 homolog B (S. cerevisiae)	25788	6.832992035	3.326538029	2.26E-14

113	CDC25A	Cell division cycle 25 homolog A (S. pombe)	993	2.754510583	1.344223745	2.06E-08
114	KIAA1524	KIAA1524	57650	10.68085201	5.256659245	1.09E-19
115	OIP5	Opa interacting protein 5	11339	4.093006024	2.028394927	4.46E-10
116	ZNF367	Zinc finger protein 367	195828	5.752741057	2.858417297	2.24E-12
117	CENPI	Centromere protein I	2491	2.828417282	1.407844865	1.24E-07
118	PLK4	Polo-like kinase 4	10733	4.461248713	2.243658574	4.72E-10
119	CKS2	CDC28 protein kinase regulatory subunit 2	1164	60.66507266	30.88699141	3.64E-24
120	BLM	Bloom syndrome, RecQ helicase-like	641	3.2371873	1.648671892	2.94E-08
121	PARPBP	PARP1 binding protein	55010	4.916851301	2.512549436	2.32E-10
122	FBXO5	F-box protein 5	26271	6.854115735	3.514321431	2.12E-13
123	CLSPN	Claspin	63967	2.665958101	1.371922225	4.58E-07
124	EZH2	Enhancer of zeste homolog 2 (Drosophila)	2146	9.854077401	5.084649728	2.44E-17
125	DSN1	DSN1, MIND kinetochore complex component, homolog (S. cerevisiae)	79980	15.94115954	8.246638055	1.08E-20
126	TCF19	Transcription factor 19	6941	10.27841898	5.322891804	1.45E-17
127	MCM2	Minichromosome maintenance complex component 2	4171	19.43072012	10.17303828	1.31E-23
128	E2F2	E2F transcription factor 2	1870	2.496417704	1.323581843	2.18E-06
129	ORC1	Origin recognition complex, subunit 1	4998	2.651307078	1.407000886	9.58E-07
130	MCM4	Minichromosome maintenance complex component 4	4173	37.95550079	20.15210467	2.65E-22
131	STIL	SCL/TAL1 interrupting locus	6491	4.625227924	2.474386004	1.61E-09

132	TIMELESS	Timeless homolog (Drosophila)	8914	19.18263837	10.27150493	1.36E-20
133	SKA1	Spindle and kinetochore associated complex subunit 1	220134	5.0268672	2.693818237	7.87E-10
134	ERCC6L	Excision repair cross-complementing rodent repair deficiency, complementation group 6-like	54821	2.502786742	1.34626481	6.01E-06
135	E2F7	E2F transcription factor 7	144455	2.443698322	1.316106371	1.06E-05
136	BRIP1	BRCA1 interacting protein C-terminal helicase 1	83990	4.231152964	2.285930939	4.56E-08
137	PCNA	Proliferating cell nuclear antigen	5111	106.8253702	57.7425208	3.24E-35
138	C15orf23	Chromosome 15 open reading frame 23	90417	12.55941208	6.83444985	1.93E-16
139	SPC25	SPC25, NDC80 kinetochore complex component, homolog (S. cerevisiae)	57405	9.525868939	5.232527087	2.57E-13
140	WDR62	WD repeat domain 62	284403	4.194362818	2.307357532	6.57E-08
141	CENPW	Centromere protein W	387103	4.102187927	2.271809473	2.91E-06
142	SKA2	Spindle and kinetochore associated complex subunit 2	348235	20.11785755	11.2532852	6.96E-17
143	H2AFZ	H2A histone family, member Z	3015	145.3840557	81.8195626	1.62E-32
144	CENPN	Centromere protein N	55839	4.773237403	2.694364663	1.12E-08
145	KNTC1	Kinetochore associated 1	9735	7.951400732	4.497711888	1.11E-10
146	CKS1B	CDC28 protein kinase regulatory subunit 1B	1163	70.44432378	39.96437789	4.41E-27

147	POLQ	Polymerase (DNA directed), theta	10721	1.843672144	1.079144425	0.000610439
148	WDHD1	WD repeat and HMG-box DNA binding protein 1	11169	4.755206652	2.789665385	3.84E-07
149	SGOL2	Shugoshin-like 2 (S. pombe)	151246	4.483234895	2.638450014	7.04E-07
150	DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta	1789	2.400776324	1.413168229	5.81E-05
151	CHEK1	Checkpoint kinase 1	1111	7.080511724	4.178165997	1.09E-09
152	MCM6	Minichromosome maintenance complex component 6	4175	19.23246548	11.36947179	1.70E-16
153	C17orf53	Chromosome 17 open reading frame 53	78995	2.677502498	1.588275804	0.000103146
154	FANCD2	Fanconi anemia, complementation group D2	2177	6.675775812	3.974084516	1.77E-08
155	C1orf135	Chromosome 1 open reading frame 135	79000	2.775367679	1.661307235	4.87E-05
156	RFC4	Replication factor C (activator 1) 4, 37kDa	5984	15.2969524	9.214279648	5.27E-14
157	LRR1	Leucine rich repeat protein 1	122769	7.48124308	4.822033333	4.78E-07
168	TUBA1C	Tubulin, alpha 1c	84790	152.748762	98.72965128	8.91E-20
159	DNAJC9	DnaJ (Hsp40) homolog, subfamily C, member 9	23234	22.43209052	14.52839977	3.23E-12
160	CCDC99	Coiled-coil domain containing 99	54908	6.785710359	4.418837584	1.75E-06
161	TUBA1B	Tubulin, alpha 1b	10376	485.5340805	317.2500722	2.31E-20
162	GMPS	Guanine monphosphate synthetase	8833	35.93670996	25.49866346	6.96E-10

Table 2: list of 19 genes that are supposed to be excellent targets to distinguish between luminal A and luminal B breast cancer

No	Gene	Description	Role in proliferation	Proliferation peak (MiSTIC)	Cellular localization
1	MCM2	Minichromosome maintenance complex component 2	Yes	Yes	Nucleus
2	CENPA	Centromere protein A	Yes	Yes	Nucleus
3	NDC80	NDC80 kinetochore complex component homolog (S. cerevisiae)	Yes	Yes	Nucleus
4	MYBL2	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	Yes	Yes	Nucleus
5	FOXM1	Forkhead box M1	Yes	Yes	Nucleus
6	E2F8	E2F transcription factor 8	Yes	Yes	Nucleus
7	E2F1	E2F transcription factor 1	Yes	Yes	Nucleus
8	UHRF1	Ubiquitin-like with PHD and ring finger domains 1	Yes	Yes	Nucleus
9	ATAD2	ATPase family, AAA domain containing 2	Yes	Yes	Nucleus
10	HMGB2	High mobility group box 2	Yes	Yes	Nucleus
11	ASF1B	ASF1 anti-silencing function 1 homolog B (S. cerevisiae)	Yes	Yes	Nucleus
12	EZH2	Enhancer of zeste homolog 2 (Drosophila)	Yes	Yes	Nucleus
13	TCF19	Transcription factor 19	Yes	Yes	Nucleus
14	E2F2	E2F transcription factor 2	Yes	Yes	Nucleus
15	MCM4	Minichromosome maintenance complex component 4	Yes	Yes	Nucleus

16	TIMELESS	Timeless homolog (Drosophila)	Yes	Yes	Nucleus
17	E2F7	E2F transcription factor 7	Yes	Yes	Nucleus
18	WDHD1	WD repeat and HMG-box DNA binding protein 1	Yes	Yes	Nucleus
19	MCM6	Minichromosome maintenance complex component 6	Yes	Yes	Nucleus

References

1. Banneau G, Guedj M, MacGrogan G, de Mascarel I, Velasco V, Schiappa R, Bonadona V, David A, Dugast C, Gilbert-Dussardier B, Ingster O, Vabres P, Caux F, de Reynies A, Iggo R, Sevenet N, Bonnet F, Longy M: **Molecular apocrine differentiation is a common feature of breast cancer in patients with germline PTEN mutations.** *Breast cancer research : BCR* 2010, **12:**R63.
2. Farmer P, Bonnefoi H, Becette V, Tubiana-Hulin M, Fumoleau P, Larsimont D, Macgrogan G, Bergh J, Cameron D, Goldstein D, Duss S, Nicoulaz AL, Brisken C, Fiche M, Delorenzi M, Iggo R: **Identification of molecular apocrine breast tumours by microarray analysis.** *Oncogene* 2005, **24:**4660-4671.
3. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS, et al: **The molecular portraits of breast tumors are conserved across microarray platforms.** *BMC genomics* 2006, **7:**96.
4. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: **Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.** *Breast cancer research : BCR* 2010, **12:**R68.
5. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100:**57-70.
6. Hanahan D, Weinberg RA: **Hallmarks of cancer: the next generation.** *Cell* 2011, **144:**646-674.
7. Bergman J: **ATP: The Perfect Energy Currency for the Cell.** *Creation Research Society Quarterly* 1999, **36**.
8. Lanigan F, O'Connor D, Martin F, Gallagher WM: **Molecular links between mammary gland development and breast cancer.** *Cellular and molecular life sciences : CMLS* 2007, **64:**3159-3184.
9. Drake R, Vogl W, Mitchell AWM: *Gray's Anatomy for Students.* 2nd edition edn: Churchill Livingstone; 2010.
10. Kumar V, Abbas AK, Aster JC: ***Robbins and Cotran Pathologic basis of disease.*** 9th edition edn: Elsevier Inc.; 2015.
11. Weigelt B, Bissell MJ: **Unraveling the microenvironmental influences on the normal mammary gland and breast cancer.** *Seminars in cancer biology* 2008, **18:**311-321.
12. Jones C, Mackay A, Grigoriadis A, Cossu A, Reis-Filho JS, Fulford L, Dexter T, Davies S, Bulmer K, Ford E, Parry S, Budroni M, Palmieri G, Neville AM, O'Hare MJ, Lakhani SR: **Expression profiling of purified normal human luminal and myoepithelial breast cells: identification of novel prognostic markers for breast cancer.** *Cancer research* 2004, **64:**3037-3045.

13. Smalley M, Ashworth A: **Stem cells and breast cancer: A field in transit.** *Nature reviews Cancer* 2003, **3:**832-844.
14. Molyneux G, Geyer FC, Magnay FA, McCarthy A, Kendrick H, Natrajan R, Mackay A, Grigoriadis A, Tutt A, Ashworth A, Reis-Filho JS, Smalley MJ: **BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells.** *Cell stem cell* 2010, **7:**403-417.
15. Conklin MW, Keely PJ: **Why the stroma matters in breast cancer: insights into breast cancer patient outcomes through the examination of stromal biomarkers.** *Cell adhesion & migration* 2012, **6:**249-260.
16. Polyak K, Hu M: **Do myoepithelial cells hold the key for breast tumor progression?** *Journal of mammary gland biology and neoplasia* 2005, **10:**231-247.
17. Elenbaas B, Spirio L, Koerner F, Fleming MD, Zimonjic DB, Donaher JL, Popescu NC, Hahn WC, Weinberg RA: **Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells.** *Genes & development* 2001, **15:**50-65.
18. Dimri G, Band H, Band V: **Mammary epithelial cell transformation: insights from cell culture and mouse models.** *Breast cancer research : BCR* 2005, **7:**171-179.
19. AmericanCancerSociety: **Cancer Facts & Figures 2014.** *Atlanta: American Cancer Society* 2014, **to find on** http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf.
20. **World Health Report 2014, Lyon. International Agency for Research on Cancer.** 2014.
21. AmericanCancerSociety: **Cancer Facts & Figures 2015.** 2015.
22. Perez EA, Moreno-Aspitia A, Aubrey Thompson E, Andorfer CA: **Adjuvant therapy of triple negative breast cancer.** *Breast cancer research and treatment* 2010, **120:**285-291.
23. Geyer FC, Rodrigues DN, Weigelt B, Reis-Filho JS: **Molecular classification of estrogen receptor-positive/luminal breast cancers.** *Advances in anatomic pathology* 2012, **19:**39-53.
24. Cummings MC, Chambers R, Simpson PT, Lakhani SR: **Molecular classification of breast cancer: is it time to pack up our microscopes?** *Pathology* 2011, **43:**1-8.
25. Colombo PE, Milanezi F, Weigelt B, Reis-Filho JS: **Microarrays in the 2010s: the contribution of microarray-based gene expression profiling to breast cancer classification, prognostication and prediction.** *Breast cancer research : BCR* 2011, **13:**212.
26. Valentin MD, da Silva SD, Privat M, Alaoui-Jamali M, Bignon YJ: **Molecular insights on basal-like breast cancer.** *Breast cancer research and treatment* 2012, **134:**21-30.
27. Trop I, LeBlanc SM, David J, Lalonde L, Tran-Thanh D, Labelle M, El Khoury MM: **Molecular classification of infiltrating breast cancer: toward**

personalized therapy. *Radiographics : a review publication of the Radiological Society of North America, Inc* 2014, **34:**1178-1195.

28. Bartlett JM, Nofech-Moses S, Rakovitch E: **Ductal carcinoma in situ of the breast: can biomarkers improve current management?** *Clinical chemistry* 2014, **60:**60-67.
29. Singhai R, Patil VW, Jaiswal SR, Patil SD, Tayade MB, Patil AV: **E-Cadherin as a diagnostic biomarker in breast cancer.** *North American journal of medical sciences* 2011, **3:**227-233.
30. Weigelt B, Geyer FC, Reis-Filho JS: **Histological types of breast cancer: how special are they?** *Molecular oncology* 2010, **4:**192-208.
31. Lakhani SR, Ellis. I.O., Schnitt, S.J., Tan, P.H., van de Vijver, M.J.: *WHO Classification of Tumours of the Breast, IARC WHO Classification of Tumours, Fourth Edition.* 2012.
32. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, Fox SB, Ichihara S, Jacquemier J, Lakhani SR, Palacios J, Richardson AL, Schnitt SJ, Schmitt FC, Tan PH, Tse GM, Badve S, Ellis IO: **Breast cancer prognostic classification in the molecular era: the role of histological grade.** *Breast cancer research : BCR* 2010, **12:**207.
33. Harbeck N, Thomssen C, Gnant M: **St. Gallen 2013: brief preliminary summary of the consensus discussion.** *Breast care* 2013, **8:**102-109.
34. **Adjuvant! Online, Decision making tools for health care professionals** [http://www.adjuvantonline.com/use.jsp]
35. Allison KH: **Molecular pathology of breast cancer: what a pathologist needs to know.** *American journal of clinical pathology* 2012, **138:**770-780.
36. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406:**747-752.
37. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, Borresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proceedings of the National Academy of Sciences of the United States of America* 2001, **98:**10869-10874.
38. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100:**8418-8423.
39. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET: **Breast cancer classification and prognosis based on gene expression profiles from a population-based study.** *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100:**10393-10398.

40. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS: **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009, **27:**1160-1167.
41. Tran B, Bedard PL: **Luminal-B breast cancer and novel therapeutic targets.** *Breast cancer research : BCR* 2011, **13:**221.
42. Knox AJ, Scaling AL, Pinto MP, Bliesner BS, Haughian JM, Abdel-Hafiz HA, Horwitz KB: **Modeling Luminal breast cancer heterogeneity: combination therapy to suppress a hormone receptor-negative, cytokeratin 5-positive subpopulation in Luminal disease.** *Breast cancer research : BCR* 2014, **16:**418.
43. Cancer Genome Atlas N: **Comprehensive molecular portraits of human breast tumours.** *Nature* 2012, **490:**61-70.
44. Prat A, Cheang MC, Martin M, Parker JS, Carrasco E, Caballero R, Tyldesley S, Gelmon K, Bernard PS, Nielsen TO, Perou CM: **Prognostic significance of progesterone receptor-positive tumor cells within immunohistochemically defined luminal A breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013, **31:**203-209.
45. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thurlimann B, Senn HJ, Panel m: **Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2013, **24:**2206-2223.
46. Guarneri V, Conte P: **Metastatic breast cancer: therapeutic options according to molecular subtypes and prior adjuvant therapy.** *The oncologist* 2009, **14:**645-656.
47. Feeley LP, Mulligan AM, Pinnaduwage D, Bull SB, Andrulis IL: **Distinguishing luminal breast cancer subtypes by Ki67, progesterone receptor or TP53 status provides prognostic information.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2014, **27:**554-561.
48. Loi S, Sotiriou C, Haibe-Kains B, Lallemand F, Conus NM, Piccart MJ, Speed TP, McArthur GA: **Gene expression profiling identifies activated growth factor signaling in poor prognosis (Luminal-B) estrogen receptor positive breast cancer.** *BMC medical genomics* 2009, **2:**37.
49. Sotiriou C, Pusztai L: **Gene-expression signatures in breast cancer.** *The New England journal of medicine* 2009, **360:**790-800.
50. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, Perou CM, Ellis MJ, Nielsen TO: **Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer.** *Journal of the National Cancer Institute* 2009, **101:**736-750.

51. Iqbal N, Iqbal N: **Human Epidermal Growth Factor Receptor 2 (HER2) in Cancers: Overexpression and Therapeutic Implications.** *Molecular biology international* 2014, **2014:**852748.
52. Gutierrez C, Schiff R: **HER2: biology, detection, and clinical implications.** *Archives of pathology & laboratory medicine* 2011, **135:**55-62.
53. Bartlett JM, Going JJ, Mallon EA, Watters AD, Reeves JR, Stanton P, Richmond J, Donald B, Ferrier R, Cooke TG: **Evaluating HER2 amplification and overexpression in breast cancer.** *The Journal of pathology* 2001, **195:**422-428.
54. Bose R, Kavuri SM, Searleman AC, Shen W, Shen D, Koboldt DC, Monsey J, Goel N, Aronson AB, Li S, Ma CX, Ding L, Mardis ER, Ellis MJ: **Activating HER2 mutations in HER2 gene amplification negative breast cancer.** *Cancer discovery* 2013, **3:**224-237.
55. Eroles P, Bosch A, Perez-Fidalgo JA, Lluch A: **Molecular biology in breast cancer: intrinsic subtypes and signaling pathways.** *Cancer treatment reviews* 2012, **38:**698-707.
56. Yersal O, Barutca S: **Biological subtypes of breast cancer: Prognostic and therapeutic implications.** *World journal of clinical oncology* 2014, **5:**412-424.
57. Gianni L, Dafni U, Gelber RD, Azambuja E, Muehlbauer S, Goldhirsch A, Untch M, Smith I, Baselga J, Jackisch C, Cameron D, Mano M, Pedrini JL, Veronesi A, Mendiola C, Pluzanska A, Semiglazov V, Vrdoljak E, Eckart MJ, Shen Z, Skiadopoulos G, Procter M, Pritchard KI, Piccart-Gebhart MJ, Bell R, Herceptin Adjuvant Trial Study T: **Treatment with trastuzumab for 1 year after adjuvant chemotherapy in patients with HER2-positive early breast cancer: a 4-year follow-up of a randomised controlled trial.** *The Lancet Oncology* 2011, **12:**236-244.
58. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L: **Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2.** *The New England journal of medicine* 2001, **344:**783-792.
59. Mittendorf EA, Wu Y, Scaltriti M, Meric-Bernstam F, Hunt KK, Dawood S, Esteva FJ, Buzdar AU, Chen H, Eksambi S, Hortobagyi GN, Baselga J, Gonzalez-Angulo AM: **Loss of HER2 amplification following trastuzumab-based neoadjuvant systemic therapy and survival outcomes.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2009, **15:**7381-7388.
60. J Baselga IB, H Eidtmann, S Di Cosimo, C Aura, E De Azambuja, H Gomez, P Dinh, K Fauria, V Van Dooren, P Paoletti, A Goldhirsch, T-W Chang, I Lang, M Untch, RD Gelber, and M Piccart-Gebhart: **First Results of the NeoALTTO Trial (BIG 01-06/EGF 106903): A Phase III, Randomized, Open Label, Neoadjuvant Study of Lapatinib, Trastuzumab, and Their Combination Plus Paclitaxel in Women with HER2-Positive Primary Breast Cancer.** *The Journal of Cancer Research* 2010, **70**.

61. L Gianni TP, Y-H Im, L Roman, L-M Tseng, M-C Liu, A Lluch-Hernandez, V Semiglazov, T Szado, and G. Ross: **Neoadjuvant Pertuzumab (P) and Trastuzumab (H): Antitumor and Safety Analysis of a Randomized Phase II Study ('NeoSphere').** *The Journal of Cancer Research* 2010, **70**.
62. Huo D, Ikpatt F, Khramtsov A, Dangou JM, Nanda R, Dignam J, Zhang B, Grushko T, Zhang C, Oluwasola O, Malaka D, Malami S, Odetunde A, Adeoye AO, Iyare F, Falusi A, Perou CM, Olopade OI: **Population differences in breast cancer: survey in indigenous African women reveals over-representation of triple-negative breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009, **27:**4515-4521.
63. Millikan RC, Newman B, Tse CK, Moorman PG, Conway K, Dressler LG, Smith LV, Labbok MH, Geradts J, Bensen JT, Jackson S, Nyante S, Livasy C, Carey L, Earp HS, Perou CM: **Epidemiology of basal-like breast cancer.** *Breast cancer research and treatment* 2008, **109:**123-139.
64. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA: **Triple-negative breast cancer: clinical features and patterns of recurrence.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2007, **13:**4429-4434.
65. Tsuda H, Takarabe T, Hasegawa F, Fukutomi T, Hirohashi S: **Large, central acellular zones indicating myoepithelial tumor differentiation in high-grade invasive ductal carcinomas as markers of predisposition to lung and brain metastases.** *The American journal of surgical pathology* 2000, **24:**197-202.
66. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, Symmans WF, Gonzalez-Angulo AM, Hennessy B, Green M, Cristofanilli M, Hortobagyi GN, Pusztai L: **Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008, **26:**1275-1281.
67. Rakha EA, El-Sayed ME, Reis-Filho J, Ellis IO: **Patho-biological aspects of basal-like breast cancer.** *Breast cancer research and treatment* 2009, **113:**411-422.
68. Rakha EA, Elsheikh SE, Aleskandarany MA, Habashi HO, Green AR, Powe DG, El-Sayed ME, Benhasouna A, Brunet JS, Akslen LA, Evans AJ, Blamey R, Reis-Filho JS, Foulkes WD, Ellis IO: **Triple-negative breast cancer: distinguishing between basal and nonbasal subtypes.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2009, **15:**2302-2310.
69. Rakha EA, Putti TC, Abd El-Rehim DM, Paish C, Green AR, Powe DG, Lee AH, Robertson JF, Ellis IO: **Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation.** *The Journal of pathology* 2006, **208:**495-506.
70. Cheang MC, Voduc D, Bajdik C, Leung S, McKinney S, Chia SK, Perou CM, Nielsen TO: **Basal-like breast cancer defined by five biomarkers has**

superior prognostic value than triple-negative phenotype. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2008, **14:**1368-1376.

71. Prat A, Adamo B, Fan C, Peg V, Vidal M, Galvan P, Vivancos A, Nuciforo P, Palmer HG, Dawood S, Rodon J, Cajal SR, Del Campo JM, Felip E, Tabernero J, Cortes J: **Genomic analyses across six cancer types identify basal-like breast cancer as a unique molecular entity.** *Scientific reports* 2013, **3:**3544.
72. van der Groep P, Bouter A, van der Zanden R, Menko FH, Buerger H, Verheijen RH, van der Wall E, van Diest PJ: **Re: Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer.** *Journal of the National Cancer Institute* 2004, **96:**712-713; author reply 714.
73. Jiang Z, Deng T, Jones R, Li H, Herschkowitz JI, Liu JC, Weigman VJ, Tsao MS, Lane TF, Perou CM, Zacksenhaus E: **Rb deletion in mouse mammary progenitors induces luminal-B or basal-like/EMT tumor subtypes depending on p53 status.** *The Journal of clinical investigation* 2010, **120:**3296-3309.
74. Liu X, Holstege H, van der Gulden H, Treur-Mulder M, Zevenhoven J, Velds A, Kerkhoven RM, van Vliet MH, Wessels LF, Peterse JL, Berns A, Jonkers J: **Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104:**12111-12116.
75. Toft DJ, Cryns VL: **Minireview: Basal-like breast cancer: from molecular profiles to targeted therapies.** *Molecular endocrinology* 2011, **25:**199-211.
76. Sorlie T: **Introducing molecular subtyping of breast cancer into the clinic?** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009, **27:**1153-1154.
77. de Ruijter TC, Veeck J, de Hoon JP, van Engeland M, Tjan-Heijnen VC: **Characteristics of triple-negative breast cancer.** *Journal of cancer research and clinical oncology* 2011, **137:**183-192.
78. Perreard L, Fan C, Quackenbush JF, Mullins M, Gauthier NP, Nelson E, Mone M, Hansen H, Buys SS, Rasmussen K, Orrico AR, Dreher D, Walters R, Parker J, Hu Z, He X, Palazzo JP, Olopade OI, Szabo A, Perou CM, Bernard PS: **Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay.** *Breast cancer research : BCR* 2006, **8:**R23.
79. Peppercorn J, Perou CM, Carey LA: **Molecular subtypes in breast cancer evaluation and management: divide and conquer.** *Cancer investigation* 2008, **26:**1-10.
80. Weigelt B, Mackay A, A'Hern R, Natrajan R, Tan DS, Dowsett M, Ashworth A, Reis-Filho JS: **Breast cancer molecular profiling with single sample predictors: a retrospective analysis.** *The Lancet Oncology* 2010, **11:**339-349.

81. Huang E, Cheng SH, Dressman H, Pittman J, Tsou MH, Horng CF, Bild A, Iversen ES, Liao M, Chen CM, West M, Nevins JR, Huang AT: **Gene expression predictors of breast cancer outcomes.** *Lancet* 2003, **361:**1590-1596.
82. Tiziana Triulzi BP, Marilena Iorio, Elda Tagliabue and Maria Luisa Carcangiu: **Insights into the heterogeneity of breast cancer: From a pathobiological to a bio-molecular classification.** *Research Signpost* 2012.
83. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C: **An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer.** *Genome biology* 2007, **8:**R157.
84. Doane AS, Danso M, Lal P, Donaton M, Zhang L, Hudis C, Gerald WL: **An estrogen receptor-negative breast cancer subset characterized by a hormonally regulated transcriptional program and response to androgen.** *Oncogene* 2006, **25:**3994-4008.
85. Schmitt FC, Reis-Filho JS: **Oncogenes, granules and breast cancer: what has c-myc to do with apocrine changes?** *Breast* 2002, **11:**463-465.
86. Lehmann-Che J, Hamy AS, Porcher R, Barritault M, Bouhidel F, Habuellelah H, Leman-Detours S, de Roquancourt A, Cahen-Doidy L, Bourstyn E, de Cremoux P, de Bazelaire C, Albiter M, Giacchetti S, Cuvier C, Janin A, Espie M, de The H, Bertheau P: **Molecular apocrine breast cancers are aggressive estrogen receptor negative tumors overexpressing either HER2 or GCDFP15.** *Breast cancer research : BCR* 2013, **15:**R37.
87. Lakis S, Kotoula V, Eleftheraki AG, Batistatou A, Bobos M, Koletsa T, Timotheadou E, Chrisafi S, Pentheroudakis G, Koutras A, Zagouri F, Linardou H, Fountzilas G: **The androgen receptor as a surrogate marker for molecular apocrine breast cancer subtyping.** *Breast* 2014, **23:**234-243.
88. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, Rasmussen KE, Jones LP, Assefnia S, Chandrasekharan S, Backlund MG, Yin Y, Khramtsov AI, Bastein R, Quackenbush J, Glazer RI, Brown PH, Green JE, Kopelovich L, Furth PA, Palazzo JP, Olopade OI, Bernard PS, Churchill GA, Van Dyke T, Perou CM: **Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors.** *Genome biology* 2007, **8:**R76.
89. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, Hartwell K, Onder TT, Gupta PB, Evans KW, Hollier BG, Ram PT, Lander ES, Rosen JM, Weinberg RA, Mani SA: **Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes.** *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107:**15449-15454.
90. Prat A, Perou CM: **Deconstructing the molecular portraits of breast cancer.** *Molecular oncology* 2011, **5:**5-23.
91. Lu S, Singh K, Mangray S, Tavares R, Noble L, Resnick MB, Yakirevich E: **Claudin expression in high-grade invasive ductal carcinoma of the breast: correlation with the molecular subtype.** *Modern pathology : an*

official journal of the United States and Canadian Academy of Pathology, Inc 2013, **26:**485-495.

92. Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, Fridlyand J, Sahin A, Agarwal R, Joy C, Liu W, Stivers D, Baggerly K, Carey M, Lluch A, Monteagudo C, He X, Weigman V, Fan C, Palazzo J, Hortobagyi GN, Nolden LK, Wang NJ, Valero V, Gray JW, Perou CM, Mills GB: **Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics.** *Cancer research* 2009, **69:**4116-4124.
93. Kothapalli R, Yoder SJ, Mane S, Loughran TP, Jr.: **Microarray results: how accurate are they?** *BMC bioinformatics* 2002, **3:**22.
94. Halgren RG, Fielden MR, Fong CJ, Zacharewski TR: **Assessment of clone identity and sequence fidelity for 1189 IMAGE cDNA clones.** *Nucleic acids research* 2001, **29:**582-588.
95. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, Pergamenschikov A, Lee JC, Lashkari D, Shalon D, Myers TG, Weinstein JN, Botstein D, Brown PO: **Systematic variation in gene expression patterns in human cancer cell lines.** *Nature genetics* 2000, **24:**227-235.
96. Taylor E, Cogdell D, Coombes K, Hu L, Ramdas L, Tabor A, Hamilton S, Zhang W: **Sequence verification as quality-control step for production of cDNA microarrays.** *BioTechniques* 2001, **31:**62-65.
97. Sotiriou C, Piccart MJ: **Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care?** *Nature reviews Cancer* 2007, **7:**545-553.
98. Tan PK, Downey TJ, Spitznagel EL, Jr., Xu P, Fu D, Dimitrov DS, Lempicki RA, Raaka BM, Cam MC: **Evaluation of gene expression measurements from commercial microarray platforms.** *Nucleic acids research* 2003, **31:**5676-5684.
99. Sinn P, Aulmann S, Wirtz R, Schott S, Marme F, Varga Z, Lebeau A, Kreipe H, Schneeweiss A: **Multigene Assays for Classification, Prognosis, and Prediction in Breast Cancer: a Critical Review on the Background and Clinical Utility.** *Geburtshilfe und Frauenheilkunde* 2013, **73:**932-940.
100. Livasy CA, Karaca G, Nanda R, Tretiakova MS, Olopade OI, Moore DT, Perou CM: **Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2006, **19:**264-271.
101. Bhargava R, Striebel J, Beriwal S, Flickinger JC, Onisko A, Ahrendt G, Dabbs DJ: **Prevalence, morphologic features and proliferation indices of breast carcinoma molecular classes using immunohistochemical surrogate markers.** *International journal of clinical and experimental pathology* 2009, **2:**444-455.
102. Won JR, Gao D, Chow C, Cheng J, Lau SY, Ellis MJ, Perou CM, Bernard PS, Nielsen TO: **A survey of immunohistochemical biomarkers for basal-like breast cancer against a gene expression profile gold standard.** *Modern*

pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 2013, **26:**1438-1450.

103. Kakimoto K, Takekoshi S, Miyajima K, Osamura RY: **Hypothesis for the mechanism for heat-induced antigen retrieval occurring on fresh frozen sections without formalin-fixation in immunohistochemistry.** *Journal of molecular histology* 2008, **39:**389-399.
104. Shi SR, Liu C, Taylor CR: **Standardization of immunohistochemistry for formalin-fixed, paraffin-embedded tissue sections based on the antigen-retrieval technique: from experiments to hypothesis.** *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* 2007, **55:**105-109.
105. O'Hurley G, Sjostedt E, Rahman A, Li B, Kampf C, Ponten F, Gallagher WM, Lindskog C: **Garbage in, garbage out: a critical evaluation of strategies used for validation of immunohistochemical biomarkers.** *Molecular oncology* 2014, **8:**783-798.
106. Warford A, Howat W, McCafferty J: **Expression profiling by high-throughput immunohistochemistry.** *Journal of immunological methods* 2004, **290:**81-92.
107. Hendriks Y, Franken P, Dierssen JW, De Leeuw W, Wijnen J, Dreef E, Tops C, Breuning M, Brocker-Vriends A, Vasen H, Fodde R, Morreau H: **Conventional and tissue microarray immunohistochemical expression analysis of mismatch repair in hereditary colorectal tumors.** *The American journal of pathology* 2003, **162:**469-477.
108. Matos LL, Trufelli DC, de Matos MG, da Silva Pinhal MA: **Immunohistochemistry as an important tool in biomarkers detection and clinical practice.** *Biomarker insights* 2010, **5:**9-20.
109. Yaziji H, Barry T: **Diagnostic Immunohistochemistry: what can go wrong?** *Advances in anatomic pathology* 2006, **13:**238-246.
110. Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, Hernandez-Boussard T, Livasy C, Cowan D, Dressler L, Akslen LA, Ragaz J, Gown AM, Gilks CB, van de Rijn M, Perou CM: **Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004, **10:**5367-5374.
111. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ, Panel m: **Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2011, **22:**1736-1747.
112. Guiu S, Michiels S, Andre F, Cortes J, Denkert C, Di Leo A, Hennessy BT, Sorlie T, Sotiriou C, Turner N, Van de Vijver M, Viale G, Loi S, Reis-Filho JS: **Molecular subclasses of breast cancer: how do we define them? The IMPAKT 2012 Working Group Statement.** *Annals of oncology : official*

journal of the European Society for Medical Oncology / ESMO 2012, **23:**2997-3006.

113. Lips EH, Mulder L, de Ronde JJ, Mandjes IA, Koolen BB, Wessels LF, Rodenhuis S, Wesseling J: **Breast cancer subtyping by immunohistochemistry and histological grade outperforms breast cancer intrinsic subtypes in predicting neoadjuvant chemotherapy response.** *Breast cancer research and treatment* 2013, **140:**63-71.
114. society Ac: **Tumor markers.** 2013.
115. Drucker E, Krapfenbauer K: **Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine.** *The EPMA journal* 2013, **4:**7.
116. Davis JC, Furstenthal L, Desai AA, Norris T, Sutaria S, Fleming E, Ma P: **The microeconomics of personalized medicine: today's challenge and tomorrow's promise.** *Nature reviews Drug discovery* 2009, **8:**279-286.
117. Ransohoff DF: **Cancer. Developing molecular biomarkers for cancer.** *Science* 2003, **299:**1679-1680.
118. Gion M, Daidone MG: **Circulating biomarkers from tumour bulk to tumour machinery: promises and pitfalls.** *European journal of cancer* 2004, **40:**2613-2622.
119. Hunter DJ, Losina E, Guermazi A, Burstein D, Lassere MN, Kraus V: **A pathway and approach to biomarker validation and qualification for osteoarthritis clinical trials.** *Current drug targets* 2010, **11:**536-545.
120. Chau CH, Rixe O, McLeod H, Figg WD: **Validation of analytic methods for biomarkers used in drug development.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2008, **14:**5967-5976.
121. Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, Keller S, Weinryb I, Green M, Duan L, Rogers JA, Millham R, O'Brien PJ, Sailstad J, Khan M, Ray C, Wagner JA: **Fit-for-purpose method development and validation for successful biomarker measurement.** *Pharmaceutical research* 2006, **23:**312-328.
122. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y: **Phases of biomarker development for early detection of cancer.** *Journal of the National Cancer Institute* 2001, **93:**1054-1061.
123. Blank M, Felice AD, Goodsaid F, Harlow P, Hausner E, Jacobson-Kram D, Taylor W, Thompson A, Throckmorton D, Xiao S: **Review of Qualification Data for Biomarkers of Nephrotoxicity Submitted by the Predictive Safety Testing Consortium.** *Center for Drug Evaluation and Research US Food and Drug Administration* 2009.
124. Goodsaid FM, Frueh FW, Mattes W: **Strategic paths for biomarker qualification.** *Toxicology* 2008, **245:**219-223.
125. Goodsaid F, Frueh F: **Biomarker qualification pilot process at the US Food and Drug Administration.** *The AAPS journal* 2007, **9:**E105-108.
126. Henry NL, Hayes DF: **Cancer biomarkers.** *Molecular oncology* 2012, **6:**140-146.

127. Pruthi S, Gostout BS, Lindor NM: **Identification and Management of Women With BRCA Mutations or Hereditary Predisposition for Breast and Ovarian Cancer.** *Mayo Clinic proceedings* 2010, **85:**1111-1120.
128. Chu CS, Rubin SC: **Screening for ovarian cancer in the general population.** *Best practice & research Clinical obstetrics & gynaecology* 2006, **20:**307-320.
129. Crawford ED, Ventii K, Shore ND: **New biomarkers in prostate cancer.** *Oncology* 2014, **28:**135-142.
130. **"Five Things Physicians and Patients Should Question"**
131. Das PM, Bast RC, Jr.: **Early detection of ovarian cancer.** *Biomarkers in medicine* 2008, **2:**291-303.
132. LaBaer J: **So, you want to look for biomarkers (introduction to the special biomarkers issue).** *Journal of proteome research* 2005, **4:**1053-1059.
133. Bialecki ES, Di Bisceglie AM: **Diagnosis of hepatocellular carcinoma.** *HPB : the official journal of the International Hepato Pancreato Biliary Association* 2005, **7:**26-34.
134. Li X, Asmitananda T, Gao L, Gai D, Song Z, Zhang Y, Ren H, Yang T, Chen T, Chen M: **Biomarkers in the lung cancer diagnosis: a clinical perspective.** *Neoplasma* 2012, **59:**500-507.
135. Weigelt B, Peterse JL, van 't Veer LJ: **Breast cancer metastasis: markers and models.** *Nature reviews Cancer* 2005, **5:**591-602.
136. Sarojini S, Tamir A, Lim H, Li S, Zhang S, Goy A, Pecora A, Suh KS: **Early detection biomarkers for ovarian cancer.** *Journal of oncology* 2012, **2012:**709049.
137. Gilligan TD, Hayes DF, Seidenfeld J, Temin S: **ASCO Clinical Practice Guideline on Uses of Serum Tumor Markers in Adult Males With Germ Cell Tumors.** *Journal of oncology practice / American Society of Clinical Oncology* 2010, **6:**199-202.
138. Bauerschlag DO, Maass N, Schem C: **Standard of care and controversies in the adjuvant endocrine treatment of hormone-responsive early breast cancer.** *Breast care* 2014, **9:**283-286.
139. Sheri A, Dowsett M: **Developments in Ki67 and other biomarkers for treatment decision making in breast cancer.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2012, **23 Suppl 10:**x219-227.
140. Salari K, Watkins H, Ashley EA: **Personalized medicine: hope or hype?** *European heart journal* 2012, **33:**1564-1570.
141. Novelli G, Ciccacci C, Borgiani P, Papaluca Amati M, Abadie E: **Genetic tests and genomic biomarkers: regulation, qualification and validation.** *Clinical cases in mineral and bone metabolism : the official journal of the Italian Society of Osteoporosis, Mineral Metabolism, and Skeletal Diseases* 2008, **5:**149-154.
142. Musolino A, Ciccolallo L, Panebianco M, Fontana E, Zanoni D, Bozzetti C, Michiara M, Silini EM, Ardizzoni A: **Multifactorial central nervous system recurrence susceptibility in patients with HER2-positive breast cancer:**

epidemiological and clinical data from a population-based cancer registry study. *Cancer* 2011, **117:**1837-1846.

143. El-Sayed MI, Maximous DW, Zakhary MM, Mikhail NN: **Biological markers and response to neoadjuvant taxane-based chemotherapy in patients with locally advanced breast cancer.** *ISRN oncology* 2012, **2012:**245891.
144. Hamilton A, Piccart M: **The contribution of molecular markers to the prediction of response in the treatment of breast cancer: a review of the literature on HER-2, p53 and BCL-2.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2000, **11:**647-663.
145. Cai FF, Kohler C, Zhang B, Chen WJ, Barekati Z, Garritsen HS, Lenner P, Toniolo P, Zhang JJ, Zhong XY: **Mutations of mitochondrial DNA as potential biomarkers in breast cancer.** *Anticancer research* 2011, **31:**4267-4271.
146. Ghaffarpour M, Mahdian R, Fereidooni F, Kamalidehghan B, Moazami N, Houshmand M: **The mitochondrial ATPase6 gene is more susceptible to mutation than the ATPase8 gene in breast cancer patients.** *Cancer cell international* 2014, **14:**21.
147. Dawson MA, Kouzarides T: **Cancer epigenetics: from mechanism to therapy.** *Cell* 2012, **150:**12-27.
148. Esteller M: **Epigenetic changes in cancer.** *F1000 biology reports* 2011, **3:**9.
149. Zhao Y, Fang X, Wang Y, Zhang J, Jiang S, Liu Z, Ma Z, Xu L, Li E, Zhang K: **Comprehensive analysis for histone acetylation of human colon cancer cells treated with a novel HDAC inhibitor.** *Current pharmaceutical design* 2014, **20:**1866-1873.
150. Silva TD, Vidigal VM, Felipe AV, JM DEL, Neto RA, Saad SS, Forones NM: **DNA methylation as an epigenetic biomarker in colorectal cancer.** *Oncology letters* 2013, **6:**1687-1692.
151. Hoque MO, Feng Q, Toure P, Dem A, Critchlow CW, Hawes SE, Wood T, Jeronimo C, Rosenbaum E, Stern J, Yu M, Trink B, Kiviat NB, Sidransky D: **Detection of aberrant methylation of four genes in plasma DNA for the detection of breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2006, **24:**4262-4269.
152. Jovanovic J, Ronneberg JA, Tost J, Kristensen V: **The epigenetics of breast cancer.** *Molecular oncology* 2010, **4:**242-254.
153. Sharma G, Mirza S, Parshad R, Gupta SD, Ralhan R: **DNA methylation of circulating DNA: a marker for monitoring efficacy of neoadjuvant chemotherapy in breast cancer patients.** *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* 2012, **33:**1837-1843.
154. Widschwendter M, Siegmund KD, Muller HM, Fiegl H, Marth C, Muller-Holzner E, Jones PA, Laird PW: **Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen.** *Cancer research* 2004, **64:**3807-3813.

155. Kiviat NB, Critchlow CW: **Novel approaches to identification of biomarkers for detection of early stage cancer.** *Disease markers* 2002, **18:**73-81.
156. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415:**530-536.
157. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, Davies SR, Snider J, Stijleman IJ, Reed J, Cheang MC, Mardis ER, Perou CM, Bernard PS, Ellis MJ: **A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2010, **16:**5222-5232.
158. Pritchard CC, Cheng HH, Tewari M: **MicroRNA profiling: approaches and considerations.** *Nature reviews Genetics* 2012, **13:**358-369.
159. Rothe F, Ignatiadis M, Chaboteaux C, Haibe-Kains B, Kheddoumi N, Majjaj S, Badran B, Fayyad-Kazan H, Desmedt C, Harris AL, Piccart M, Sotiriou C: **Global microRNA expression profiling identifies MiR-210 associated with tumor proliferation, invasion and poor clinical outcome in breast cancer.** *PloS one* 2011, **6:**e20980.
160. Camps C, Buffa FM, Colella S, Moore J, Sotiriou C, Sheldon H, Harris AL, Gleadle JM, Ragoussis J: **hsa-miR-210 Is induced by hypoxia and is an independent prognostic factor in breast cancer.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2008, **14:**1340-1348.
161. Gasparini P, Cascione L, Fassan M, Lovat F, Guler G, Balci S, Irkkan C, Morrison C, Croce CM, Shapiro CL, Huebner K: **microRNA expression profiling identifies a four microRNA signature as a novel diagnostic and prognostic biomarker in triple negative breast cancers.** *Oncotarget* 2014, **5:**1174-1184.
162. Zhou M, Liu Z, Zhao Y, Ding Y, Liu H, Xi Y, Xiong W, Li G, Lu J, Fodstad O, Riker AI, Tan M: **MicroRNA-125b confers the resistance of breast cancer cells to paclitaxel through suppression of pro-apoptotic Bcl-2 antagonist killer 1 (Bak1) expression.** *The Journal of biological chemistry* 2010, **285:**21496-21507.
163. Hiroshi Ichikawaa b, Tatsuo Kandac, Shin-ichi Kosugib, Yasuyuki Kawachid, Toshifumi Wakaib, Tadashi Kondoa: **Proteomic and meta-transcriptomic study on lymph node metastasis in gastric cancer.** *EuPA Open Proteomics* 2014, **3:**183-194.
164. Uemura N, Kondo T: **Current status of predictive biomarkers for neoadjuvant therapy in esophageal cancer.** *World journal of gastrointestinal pathophysiology* 2014, **5:**322-334.
165. Mishra A, Verma M: **Cancer biomarkers: are we ready for the prime time?** *Cancers* 2010, **2:**190-208.

166. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, Somerfield MR, Hayes DF, Bast RC, Jr., American Society of Clinical O: **American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007, **25:**5287-5312.
167. Sturgeon CM, Duffy MJ, Stenman UH, Lilja H, Brunner N, Chan DW, Babaian R, Bast RC, Jr., Dowell B, Esteva FJ, Haglund C, Harbeck N, Hayes DF, Holten-Andersen M, Klee GG, Lamerz R, Looijenga LH, Molina R, Nielsen HJ, Rittenhouse H, Semjonow A, Shih Ie M, Sibley P, Soletormos G, Stephan C, Sokoll L, Hoffman BR, Diamandis EP, National Academy of Clinical B: **National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers.** *Clinical chemistry* 2008, **54:**e11-79.
168. Lee JW, Figeys D, Vasilescu J: **Biomarker assay translation from discovery to clinical studies in cancer drug development: quantification of emerging protein biomarkers.** *Advances in cancer research* 2007, **96:**269-298.
169. Issaq HJ, Waybright TJ, Veenstra TD: **Cancer biomarker discovery: Opportunities and pitfalls in analytical methods.** *Electrophoresis* 2011, **32:**967-975.
170. Baggerly KA, Morris JS, Coombes KR: **Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments.** *Bioinformatics* 2004, **20:**777-785.
171. Eric A. Libré HM, Alexander I. Spira: **Let's test: A Multidisciplinary Approach to Biomarker Testing in NSCLC.** *Boehringer Ingelheim Pharmaceuticals* 2013.
172. Moore HM, Compton CC, Lim MD, Vaught J, Christiansen KN, Alper J: **2009 Biospecimen research network symposium: advancing cancer research through biospecimen science.** *Cancer research* 2009, **69:**6770-6772.
173. Ransohoff DF, Gourlay ML: **Sources of bias in specimens for research about molecular markers for cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010, **28:**698-704.
174. Zolg W: **The proteomic search for diagnostic biomarkers: lost in translation?** *Molecular & cellular proteomics : MCP* 2006, **5:**1720-1726.
175. Poste G: **Bring on the biomarkers.** *Nature* 2011, **469:**156-157.
176. Elias JE, Haas W, Faherty BK, Gygi SP: **Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations.** *Nature methods* 2005, **2:**667-675.
177. Garnett MJ, Edelman EJ, Heidorn SJ, Greenman CD, Dastur A, Lau KW, Greninger P, Thompson IR, Luo X, Soares J, Liu Q, Iorio F, Surdez D, Chen L, Milano RJ, Bignell GR, Tam AT, Davies H, Stevenson JA, Barthorpe S, Lutz SR, Kogera F, Lawrence K, McLaren-Douglas A, Mitropoulos X, Mironenko T, Thi H, Richardson L, Zhou W, Jewitt F, et al: **Systematic identification of**

genomic markers of drug sensitivity in cancer cells. *Nature* 2012, **483:**570-575.

178. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, Reddy A, Liu M, Murray L, Berger MF, Monahan JE, Morais P, Meltzer J, Korejwa A, Jane-Valbuena J, Mapa FA, Thibault J, Bric-Furlong E, Raman P, Shipway A, Engels IH, Cheng J, Yu GK, Yu J, Aspesi P, Jr., de Silva M, et al: **The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.** *Nature* 2012, **483:**603-607.
179. Burdall SE, Hanby AM, Lansdown MR, Speirs V: **Breast cancer cell lines: friend or foe?** *Breast cancer research : BCR* 2003, **5:**89-95.
180. **Cancer Research Tools for Biological Relevance** [http://bio.lonza.com/uploads/tx_mwaxmarketingmaterial/Lonza_Brochures_Cancer_Research_Tools_for_Biological_Relevance.pdf]
181. Meric-Bernstam F, Mills GB: **Overcoming implementation challenges of personalized cancer therapy.** *Nature reviews Clinical oncology* 2012, **9:**542-548.
182. Anderson NL, Polanski M, Pieper R, Gatlin T, Tirumalai RS, Conrads TP, Veenstra TD, Adkins JN, Pounds JG, Fagan R, Lobley A: **The human plasma proteome: a nonredundant list developed by combination of four separate sources.** *Molecular & cellular proteomics : MCP* 2004, **3:**311-326.
183. Buscher JM, Czernik D, Ewald JC, Sauer U, Zamboni N: **Cross-platform comparison of methods for quantitative metabolomics of primary metabolism.** *Analytical chemistry* 2009, **81:**2135-2143.
184. Hudler P, Kocevar N, Komel R: **Proteomic approaches in biomarker discovery: new perspectives in cancer diagnostics.** *TheScientificWorldJournal* 2014, **2014:**260348.
185. Eric Groves JH, Christopher Ung: **From Biomarkers to Diagnostics – The Road to Success.** In *Book From Biomarkers to Diagnostics – The Road to Success* (Editor ed.^eds.). City; 2010.
186. Reis-Filho JS, Pusztai L: **Gene expression profiling in breast cancer: classification, prognostication, and prediction.** *Lancet* 2011, **378:**1812-1823.
187. Reis-Filho JS, Weigelt B, Fumagalli D, Sotiriou C: **Molecular profiling: moving away from tumor philately.** *Science translational medicine* 2010, **2:**47ps43.
188. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, Desmedt C, Ignatiadis M, Sengstag T, Schutz F, Goldstein DR, Piccart M, Delorenzi M: **Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures.** *Breast cancer research : BCR* 2008, **10:**R65.
189. George SL: **Statistical issues in translational cancer research.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2008, **14:**5954-5958.

190. Poste G: **Biospecimens, biomarkers, and burgeoning data: the imperative for more rigorous research standards.** *Trends in molecular medicine* 2012, **18:**717-722.
191. Marino M, Galluzzo P, Ascenzi P: **Estrogen signaling multiple pathways to impact gene transcription.** *Current genomics* 2006, **7:**497-508.
192. Herynk MH, Fuqua SA: **Estrogen receptor mutations in human disease.** *Endocrine reviews* 2004, **25:**869-898.
193. Liang J, Shang Y: **Estrogen and cancer.** *Annual review of physiology* 2013, **75:**225-240.
194. Kuiper GG, Carlsson B, Grandien K, Enmark E, Haggblad J, Nilsson S, Gustafsson JA: **Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta.** *Endocrinology* 1997, **138:**863-870.
195. Murillo-Ortiz B, Perez-Luque E, Malacara JM, Daza-Benitez L, Hernandez-Gonzalez M, Benitez-Bribiesca L: **Expression of estrogen receptor alpha and beta in breast cancers of pre- and post-menopausal women.** *Pathology oncology research : POR* 2008, **14:**435-442.
196. Matthews J, Gustafsson JA: **Estrogen signaling: a subtle balance between ER alpha and ER beta.** *Molecular interventions* 2003, **3:**281-292.
197. Thomas C, Gustafsson JA: **The different roles of ER subtypes in cancer biology and therapy.** *Nature reviews Cancer* 2011, **11:**597-608.
198. Tiano JP, Mauvais-Jarvis F: **Importance of oestrogen receptors to preserve functional beta-cell mass in diabetes.** *Nature reviews Endocrinology* 2012, **8:**342-351.
199. Heldring N, Pike A, Andersson S, Matthews J, Cheng G, Hartman J, Tujague M, Strom A, Treuter E, Warner M, Gustafsson JA: **Estrogen receptors: how do they signal and what are their targets.** *Physiological reviews* 2007, **87:**905-931.
200. Wong RL, Walker CL: **Molecular pathways: environmental estrogens activate nongenomic signaling to developmentally reprogram the epigenome.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2013, **19:**3732-3737.
201. Bjornstrom L, Sjoberg M: **Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes.** *Molecular endocrinology* 2005, **19:**833-842.
202. Umekita Y, Souda M, Ohi Y, Rai Y, Sagara Y, Sagara Y, Yoshida H: **Expression of estrogen receptor alpha and progesterone receptor in normal human breast epithelium.** *In vivo* 2007, **21:**535-539.
203. Lawson JS, Field AS, Tran DD, Killeen J, Maskarenic G, Ishikura H, Trichopoulos D: **Breast cancer incidence and estrogen receptor alpha in normal mammary tissue--an epidemiologic study among Japanese women in Japan and Hawaii.** *International journal of cancer Journal international du cancer* 2002, **97:**685-687.

204. Khan SA, Rogers MA, Khurana KK, Meguid MM, Numann PJ: **Estrogen receptor expression in benign breast epithelium and breast cancer risk.** *Journal of the National Cancer Institute* 1998, **90:**37-42.
205. Khan SA, Yee KA, Kaplan C, Siddiqui JF: **Estrogen receptor alpha expression in normal human breast epithelium is consistent over time.** *International journal of cancer Journal international du cancer* 2002, **102:**334-337.
206. Uchida N, Suda T, Ishiguro K: **Effect of chemotherapy for luminal a breast cancer.** *Yonago acta medica* 2013, **56:**51-56.
207. Qureshi A, Pervez S: **Allred scoring for ER reporting and it's impact in clearly distinguishing ER negative from ER positive breast cancers.** *JPMA The Journal of the Pakistan Medical Association* 2010, **60:**350-353.
208. Allred DC, Harvey JM, Berardo M, Clark GM: **Prognostic and predictive factors in breast cancer by immunohistochemical analysis.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 1998, **11:**155-168.
209. Allred C: **Assessment of Prognostic and Predictive Factors in Breast Cancer by Immunohistochemistry** http://www.dako.com/28824_2006_conn9_prognostic_factors_allred.pdf 2006.
210. Harvey JM, Clark GM, Osborne CK, Allred DC: **Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 1999, **17:**1474-1481.
211. Early Breast Cancer Trialists' Collaborative G, Clarke M, Coates AS, Darby SC, Davies C, Gelber RD, Godwin J, Goldhirsch A, Gray R, Peto R, Pritchard KI, Wood WC: **Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials.** *Lancet* 2008, **371:**29-40.
212. Garcia-Becerra R, Santos N, Diaz L, Camacho J: **Mechanisms of Resistance to Endocrine Therapy in Breast Cancer: Focus on Signaling Pathways, miRNAs and Genetically Based Resistance.** *International journal of molecular sciences* 2012, **14:**108-145.
213. Chung YL, Sheu ML, Yang SC, Lin CH, Yen SH: **Resistance to tamoxifen-induced apoptosis is associated with direct interaction between Her2/neu and cell membrane estrogen receptor in breast cancer.** *International journal of cancer Journal international du cancer* 2002, **97:**306-312.
214. Sabnis G, Schayowitz A, Goloubeva O, Macedo L, Brodie A: **Trastuzumab reverses letrozole resistance and amplifies the sensitivity of breast cancer cells to estrogen.** *Cancer research* 2009, **69:**1416-1428.
215. Paplomata E, O'Regan R: **New and emerging treatments for estrogen receptor-positive breast cancer: focus on everolimus.** *Therapeutics and clinical risk management* 2013, **9:**27-36.

216. Shoman N, Klassen S, McFadden A, Bickis MG, Torlakovic E, Chibbar R: **Reduced PTEN expression predicts relapse in patients with breast carcinoma treated by tamoxifen.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2005, **18:**250-259.
217. Ades F, Zardavas D, Bozovic-Spasojevic I, Pugliano L, Fumagalli D, de Azambuja E, Viale G, Sotiriou C, Piccart M: **Luminal B breast cancer: molecular characterization, clinical management, and future perspectives.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2014, **32:**2794-2803.
218. Hagan CR, Daniel AR, Dressing GE, Lange CA: **Role of phosphorylation in progesterone receptor signaling and specificity.** *Molecular and cellular endocrinology* 2012, **357:**43-49.
219. Gadkar-Sable S, Shah C, Rosario G, Sachdeva G, Puri C: **Progesterone receptors: various forms and functions in reproductive tissues.** *Frontiers in bioscience : a journal and virtual library* 2005, **10:**2118-2130.
220. Wei LL, Norris BM, Baker CJ: **An N-terminally truncated third progesterone receptor protein, PR(C), forms heterodimers with PR(B) but interferes in PR(B)-DNA binding.** *The Journal of steroid biochemistry and molecular biology* 1997, **62:**287-297.
221. Law ML, Kao FT, Wei Q, Hartz JA, Greene GL, Zarucki-Schulz T, Conneely OM, Jones C, Puck TT, O'Malley BW, et al.: **The progesterone receptor gene maps to human chromosome band 11q13, the site of the mammary oncogene int-2.** *Proceedings of the National Academy of Sciences of the United States of America* 1987, **84:**2877-2881.
222. Wetendorf M, DeMayo FJ: **Progesterone receptor signaling in the initiation of pregnancy and preservation of a healthy uterus.** *The International journal of developmental biology* 2014, **58:**95-106.
223. Kim JJ, Kurita T, Bulun SE: **Progesterone action in endometrial cancer, endometriosis, uterine fibroids, and breast cancer.** *Endocrine reviews* 2013, **34:**130-162.
224. Bland R: **Steroid hormone receptor expression and action in bone.** *Clinical science* 2000, **98:**217-240.
225. Nakamura Y, Suzuki T, Inoue T, Tazawa C, Ono K, Moriya T, Saito H, Ishibashi T, Takahashi S, Yamada S, Sasano H: **Progesterone receptor subtypes in vascular smooth muscle cells of human aorta.** *Endocrine journal* 2005, **52:**245-252.
226. German-Castelan L, Manjarrez-Marmolejo J, Gonzalez-Arenas A, Gonzalez-Moran MG, Camacho-Arroyo I: **Progesterone induces the growth and infiltration of human astrocytoma cells implanted in the cerebral cortex of the rat.** *Biomed Res Int* 2014, **2014:**393174.
227. Scarpin KM, Graham JD, Mote PA, Clarke CL: **Progesterone action in human tissues: regulation by progesterone receptor (PR) isoform expression, nuclear positioning and coregulator expression.** *Nuclear receptor signaling* 2009, **7:**e009.

228. Bartlett JM, Brookes CL, Robson T, van de Velde CJ, Billingham LJ, Campbell FM, Grant M, Hasenburg A, Hille ET, Kay C, Kieback DG, Putter H, Markopoulos C, Kranenbarg EM, Mallon EA, Dirix L, Seynaeve C, Rea D: **Estrogen receptor and progesterone receptor as predictive biomarkers of response to endocrine therapy: a prospectively powered pathology study in the Tamoxifen and Exemestane Adjuvant Multinational trial.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011, **29:**1531-1538.
229. Tovey S, Dunne B, Witton CJ, Forsyth A, Cooke TG, Bartlett JM: **Can molecular markers predict when to implement treatment with aromatase inhibitors in invasive breast cancer?** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2005, **11:**4835-4842.
230. Bardou VJ, Arpino G, Elledge RM, Osborne CK, Clark GM: **Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2003, **21:**1973-1979.
231. Liu S, Chia SK, Mehl E, Leung S, Rajput A, Cheang MC, Nielsen TO: **Progesterone receptor is a significant factor associated with clinical outcomes and effect of adjuvant tamoxifen therapy in breast cancer patients.** *Breast cancer research and treatment* 2010, **119:**53-61.
232. Stendahl M, Ryden L, Nordenskjold B, Jonsson PE, Landberg G, Jirstrom K: **High progesterone receptor expression correlates to the effect of adjuvant tamoxifen in premenopausal breast cancer patients.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2006, **12:**4614-4618.
233. Cancello G, Maisonneuve P, Rotmensz N, Viale G, Mastropasqua MG, Pruneri G, Montagna E, Iorfida M, Mazza M, Balduzzi A, Veronesi P, Luini A, Intra M, Goldhirsch A, Colleoni M: **Progesterone receptor loss identifies Luminal B breast cancer subgroups at higher risk of relapse.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2013, **24:**661-668.
234. Dunnwald LK, Rossing MA, Li CI: **Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients.** *Breast cancer research : BCR* 2007, **9:**R6.
235. Dowsett M, Cuzick J, Wale C, Howell T, Houghton J, Baum M: **Retrospective analysis of time to recurrence in the ATAC trial according to hormone receptor status: an hypothesis-generating study.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005, **23:**7512-7517.
236. Hefti MM, Hu R, Knoblauch NW, Collins LC, Haibe-Kains B, Tamimi RM, Beck AH: **Estrogen receptor negative/progesterone receptor positive breast cancer is not a reproducible subtype.** *Breast cancer research : BCR* 2013, **15:**R68.

237. Tadashi Yamamoto MS, Kentaro Kumazawa, Ayano Doi, Atsuka Matsui, Shiori Takebe, Takuya Amari, Masaaki Oyama and Kentaro Semba: **ErbB2/HER2: Its Contribution to Basic Cancer Biology and the Development of Molecular Targeted Therapy.** In *Breast Cancer - Carcinogenesis, Cell Growth and Signalling Pathways.* Edited by Gunduz M. Japan: InTech; 2011
238. Sliwkowski MX: **Ready to partner.** *Nature structural biology* 2003, **10:**158-159.
239. Ghosh R, Narasanna A, Wang SE, Liu S, Chakrabarty A, Balko JM, Gonzalez-Angulo AM, Mills GB, Penuel E, Winslow J, Sperinde J, Dua R, Pidaparthi S, Mukherjee A, Leitzel K, Kostler WJ, Lipton A, Bates M, Arteaga CL: **Trastuzumab has preferential activity against breast cancers driven by HER2 homodimers.** *Cancer research* 2011, **71:**1871-1882.
240. Nahta R, Yuan LX, Zhang B, Kobayashi R, Esteva FJ: **Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells.** *Cancer research* 2005, **65:**11118-11128.
241. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu PB, Paik S, Perez EA, Press MF, Spears PA, Vance GH, Viale G, Hayes DF, American Society of Clinical O, College of American P: **Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update.** *Archives of pathology & laboratory medicine* 2014, **138:**241-256.
242. Vici P, Pizzuti L, Natoli C, Gamucci T, Di Lauro L, Barba M, Sergi D, Botti C, Michelotti A, Moscetti L, Mariani L, Izzo F, D'Onofrio L, Sperduti I, Conti F, Rossi V, Cassano A, Maugeri-Sacca M, Mottolese M, Marchetti P: **Triple positive breast cancer: A distinct subtype?** *Cancer treatment reviews* 2015, **41:**69-76.
243. Cianfrocca M, Goldstein LJ: **Prognostic and predictive factors in early-stage breast cancer.** *The oncologist* 2004, **9:**606-616.
244. Borg A, Tandon AK, Sigurdsson H, Clark GM, Ferno M, Fuqua SA, Killander D, McGuire WL: **HER-2/neu amplification predicts poor survival in node-positive breast cancer.** *Cancer research* 1990, **50:**4332-4337.
245. Paterson MC, Dietrich KD, Danyluk J, Paterson AH, Lees AW, Jamil N, Hanson J, Jenkins H, Krause BE, McBlain WA, et al.: **Correlation between c-erbB-2 amplification and risk of recurrent disease in node-negative breast cancer.** *Cancer research* 1991, **51:**556-567.
246. Winstanley J, Cooke T, Murray GD, Platt-Higgins A, George WD, Holt S, Myskov M, Spedding A, Barraclough BR, Rudland PS: **The long term prognostic significance of c-erbB-2 in primary breast cancer.** *British journal of cancer* 1991, **63:**447-450.
247. Moon YW, Park S, Sohn JH, Kang DR, Koo JS, Park HS, Chung HC, Park BW: **Clinical significance of progesterone receptor and HER2 status in estrogen receptor-positive, operable breast cancer with adjuvant**

tamoxifen. *Journal of cancer research and clinical oncology* 2011, **137:**1123-1130.

248. Ellis MJ, Coop A, Singh B, Mauriac L, Llombert-Cussac A, Janicke F, Miller WR, Evans DB, Dugan M, Brady C, Quebe-Fehling E, Borgs M: **Letrozole is more effective neoadjuvant endocrine therapy than tamoxifen for ErbB-1- and/or ErbB-2-positive, estrogen receptor-positive primary breast cancer: evidence from a phase III randomized trial.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2001, **19:**3808-3816.
249. Borg A, Baldetorp B, Ferno M, Killander D, Olsson H, Ryden S, Sigurdsson H: **ERBB2 amplification is associated with tamoxifen resistance in steroid-receptor positive breast cancer.** *Cancer letters* 1994, **81:**137-144.
250. Carlomagno C, Perrone F, Gallo C, De Laurentiis M, Lauria R, Morabito A, Pettinato G, Panico L, D'Antonio A, Bianco AR, De Placido S: **c-erb B2 overexpression decreases the benefit of adjuvant tamoxifen in early-stage breast cancer without axillary lymph node metastases.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 1996, **14:**2702-2708.
251. De Laurentiis M, Arpino G, Massarelli E, Ruggiero A, Carlomagno C, Ciardiello F, Tortora G, D'Agostino D, Caputo F, Cancello G, Montagna E, Malorni L, Zinno L, Lauria R, Bianco AR, De Placido S: **A meta-analysis on the interaction between HER-2 expression and response to endocrine treatment in advanced breast cancer.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2005, **11:**4741-4748.
252. Bozzetti C, Musolino A, Camisa R, Bisagni G, Flora M, Bassano C, Martella E, Lagrasta C, Nizzoli R, Personeni N, Leonardi F, Cocconi G, Ardizzoni A: **Evaluation of HER-2/neu amplification and other biological markers as predictors of response to neoadjuvant anthracycline-based chemotherapy in primary breast cancer: the role of anthracycline dose intensity.** *American journal of clinical oncology* 2006, **29:**171-177.
253. Learn PA, Yeh IT, McNutt M, Chisholm GB, Pollock BH, Rousseau DL, Jr., Sharkey FE, Cruz AB, Kahlenberg MS: **HER-2/neu expression as a predictor of response to neoadjuvant docetaxel in patients with operable breast carcinoma.** *Cancer* 2005, **103:**2252-2260.
254. Kaufmann M, von Minckwitz G, Bear HD, Buzdar A, McGale P, Bonnefoi H, Colleoni M, Denkert C, Eiermann W, Jackesz R, Makris A, Miller W, Pierga JY, Semiglazov V, Schneeweiss A, Souchon R, Stearns V, Untch M, Loibl S: **Recommendations from an international expert panel on the use of neoadjuvant (primary) systemic treatment of operable breast cancer: new perspectives 2006.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2007, **18:**1927-1934.
255. Guiu S, Mouret Reynier MA, Toure M, Coudert B: **Predictive Factors of Response in HER2-Positive Breast Cancer Treated by Neoadjuvant Therapy.** *Journal of oncology* 2013, **2013:**854121.

256. Dawood S, Broglio K, Buzdar AU, Hortobagyi GN, Giordano SH: **Prognosis of women with metastatic breast cancer by HER2 status and trastuzumab treatment: an institutional-based review.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010, **28:**92-98.
257. Albain KS, Paik S, van't Veer L: **Prediction of adjuvant chemotherapy benefit in endocrine responsive, early breast cancer using multigene assays.** *Breast* 2009, **18 Suppl 3:**S141-145.
258. Engelhardt EG, Garvelink MM, de Haes JH, van der Hoeven JJ, Smets EM, Pieterse AH, Stiggelbout AM: **Predicting and communicating the risk of recurrence and death in women with early-stage breast cancer: a systematic review of risk prediction models.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2014, **32:**238-250.
259. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.** *The New England journal of medicine* 2004, **351:**2817-2826.
260. Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, Baker J, Walker M, Watson D, Hackett J, Blick NT, Greenberg D, Fehrenbacher L, Langholz B, Quesenberry CP: **A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients.** *Breast cancer research : BCR* 2006, **8:**R25.
261. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE, Jr., Wickerham DL, Wolmark N: **Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2006, **24:**3726-3734.
262. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF, Breast Cancer Intergroup of North A: **Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial.** *The Lancet Oncology* 2010, **11:**55-65.
263. Goldstein LJ, Gray R, Badve S, Childs BH, Yoshizawa C, Rowley S, Shak S, Baehner FL, Ravdin PM, Davidson NE, Sledge GW, Jr., Perez EA, Shulman LN, Martino S, Sparano JA: **Prognostic utility of the 21-gene assay in hormone receptor-positive operable breast cancer compared with classical clinicopathologic features.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008, **26:**4063-4071.
264. Oratz R, Kim B, Chao C, Skrzypczak S, Ory C, Bugarini R, Broder M: **Physician survey of the effect of the 21-gene recurrence score assay results on treatment recommendations for patients with lymph node-positive,**

estrogen receptor-positive breast cancer. *Journal of oncology practice / American Society of Clinical Oncology* 2011, **7:**94-99.

265. Mamounas EP, Tang G, Fisher B, Paik S, Shak S, Costantino JP, Watson D, Geyer CE, Jr., Wickerham DL, Wolmark N: **Association between the 21-gene recurrence score assay and risk of locoregional recurrence in node-negative, estrogen receptor-positive breast cancer: results from NSABP B-14 and NSABP B-20.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2010, **28:**1677-1683.
266. Zujewski JA, Kamin L: **Trial assessing individualized options for treatment for breast cancer: the TAILORx trial.** *Future oncology* 2008, **4:**603-610.
267. Paik S: **Development and clinical utility of a 21-gene recurrence score prognostic assay in patients with early breast cancer treated with tamoxifen.** *The oncologist* 2007, **12:**631-635.
268. **NCCN clinical practice guidelines in oncology (NCCN guidelines) Breast cancer.** http://wwwnccnorg/patients/ Version 3.2014.
269. **St. Gallen International Breast Cancer Expert Panel Guidelines Include Oncotype DX(R) as Predictor of Chemotherapy Benefit.** *Geneva, Switzerland and Redwood City, CA: Genomic Health, Inc; July 25, 2011* 2011.
270. Ingoldsby HM: **IDENTIFICATION OF PROGNOSTIC TARGETS IN BREAST CANCER USING IMMUNOHISTOCHEMICAL AND IN SITU HYBRIDISATION METHODOLOGIES.** National University of Ireland, Department of pathology; 2011.
271. **PAM50-Based Prosigna Breast Cancer Assay Helps to Identify Patients at Risk of Late Distant Recurrence in a Combined Analysis of 2,137 Patients. Prosigna Combined Analysis of TransATAC and ABCGS-8 Data Presented at the 2013 San Antonio Breast Cancer Symposium. it will be found at** http://www.nanostring.com/company/corp_press_release?id=100
272. Tian S, Roepman P, Van't Veer LJ, Bernards R, de Snoo F, Glas AM: **Biological functions of the genes in the mammaprint breast cancer profile reflect the hallmarks of cancer.** *Biomarker insights* 2010, **5:**129-138.
273. Kittaneh M, Montero AJ, Gluck S: **Molecular profiling for breast cancer: a comprehensive review.** *Biomarkers in cancer* 2013, **5:**61-70.
274. Knauer M, Mook S, Rutgers EJ, Bender RA, Hauptmann M, van de Vijver MJ, Koornstra RH, Bueno-de-Mesquita JM, Linn SC, van 't Veer LJ: **The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer.** *Breast cancer research and treatment* 2010, **120:**655-661.
275. E. Rutgers P-GM, Bogaerts J, Delaloge S, Van ‘t Veer LJ, Rubio IT, Viale G, Nitz U, Pierga JY, Vindevoghel A, Brain E, Ravdin PM, Messina C, Cardoso F on behalf of the MINDACT TRANSBIG studygroup. : **Baseline results of the EORTC 10041/MINDACT TRIAL (Microarray In Node 0-3 positive Disease may Avoid ChemoTherapy). Presented during the “Breast Cancer – Early Disease”** 2013, **Poster Session on Monday 30 September 2013**.

276. Bastien RR, Rodriguez-Lescure A, Ebbert MT, Prat A, Munarriz B, Rowe L, Miller P, Ruiz-Borrego M, Anderson D, Lyons B, Alvarez I, Dowell T, Wall D, Segui MA, Barley L, Boucher KM, Alba E, Pappas L, Davis CA, Aranda I, Fauron C, Stijleman IJ, Palacios J, Anton A, Carrasco E, Caballero R, Ellis MJ, Nielsen TO, Perou CM, Astill M, et al: **PAM50 breast cancer subtyping by RT-qPCR and concordance with standard clinical molecular markers.** *BMC medical genomics* 2012, **5:**44.
277. Kelly CM, Bernard PS, Krishnamurthy S, Wang B, Ebbert MT, Bastien RR, Boucher KM, Young E, Iwamoto T, Pusztai L: **Agreement in risk prediction between the 21-gene recurrence score assay (Oncotype DX(R)) and the PAM50 breast cancer intrinsic Classifier in early-stage estrogen receptor-positive breast cancer.** *The oncologist* 2012, **17:**492-498.
278. Prat A, Parker JS, Fan C, Cheang MC, Miller LD, Bergh J, Chia SK, Bernard PS, Nielsen TO, Ellis MJ, Carey LA, Perou CM: **Concordance among gene expression-based predictors for ER-positive breast cancer treated with adjuvant tamoxifen.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2012, **23:**2866-2873.
279. Nielsen T, Wallden B, Schaper C, Ferree S, Liu S, Gao D, Barry G, Dowidar N, Maysuria M, Storhoff J: **Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens.** *BMC cancer* 2014, **14:**177.
280. Aleskandarany MA, Green AR, Benhasouna AA, Barros FF, Neal K, Reis-Filho JS, Ellis IO, Rakha EA: **Prognostic value of proliferation assay in the luminal, HER2-positive, and triple-negative biologic classes of breast cancer.** *Breast cancer research : BCR* 2012, **14:**R3.
281. Jonat W, Arnold N: **Is the Ki-67 labelling index ready for clinical use?** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2011, **22:**500-502.
282. Veronese SM, Gambacorta M, Gottardi O, Scanzi F, Ferrari M, Lampertico P: **Proliferation index as a prognostic marker in breast cancer.** *Cancer* 1993, **71:**3926-3931.
283. Reena RM, Mastura M, Siti-Aishah MA, Munirah MA, Norlia A, Naqiyah I, Rohaizak M, Sharifah NA: **Minichromosome maintenance protein 2 is a reliable proliferative marker in breast carcinoma.** *Annals of diagnostic pathology* 2008, **12:**340-343.
284. Lee LH, Yang H, Bigras G: **Current breast cancer proliferative markers correlate variably based on decoupled duration of cell cycle phases.** *Scientific reports* 2014, **4:**5122.
285. Elston CW, Ellis IO: **Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.** *Histopathology* 2002, **41:**154-161.
286. Trihia H, Murray S, Price K, Gelber RD, Golouh R, Goldhirsch A, Coates AS, Collins J, Castiglione-Gertsch M, Gusterson BA, International Breast Cancer Study G: **Ki-67 expression in breast carcinoma: its association with**

grading systems, clinical parameters, and other prognostic factors--a surrogate marker? *Cancer* 2003, **97:**1321-1331.

287. Medri L, Volpi A, Nanni O, Vecci AM, Mangia A, Schittulli F, Padovani F, Giunchi DC, Zito A, Amadori D, Paradiso A, Silvestrini R: **Prognostic relevance of mitotic activity in patients with node-negative breast cancer.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2003, **16:**1067-1075.
288. Chowdhury N, Pai MR, Lobo FD, Kini H, Varghese R: **Interobserver variation in breast cancer grading: a statistical modeling approach.** *Analytical and quantitative cytology and histology / the International Academy of Cytology [and] American Society of Cytology* 2006, **28:**213-218.
289. Elston CW: **Classification and grading of invasive breast carcinoma.** *Verhandlungen der Deutschen Gesellschaft fur Pathologie* 2005, **89:**35-44.
290. Genestie C, Zafrani B, Asselain B, Fourquet A, Rozan S, Validire P, Vincent-Salomon A, Sastre-Garau X: **Comparison of the prognostic value of Scarff-Bloom-Richardson and Nottingham histological grades in a series of 825 cases of breast cancer: major importance of the mitotic count as a component of both grading systems.** *Anticancer research* 1998, **18:**571-576.
291. Biesterfeld S, Noll I, Noll E, Wohltmann D, Bocking A: **Mitotic frequency as a prognostic factor in breast cancer.** *Human pathology* 1995, **26:**47-52.
292. Tryfonidis K, Kafousi M, Perraki M, Apostolaki S, Agelaki S, Georgoulias V, Stathopoulos E, Mavroudis D: **Detection of circulating cytokeratin-19 mRNA-positive cells in the blood and the mitotic index of the primary tumor have independent prognostic value in early breast cancer.** *Clinical breast cancer* 2014, **14:**442-450.
293. Laroye GJ, Minkin S: **The impact of mitotic index on predicting outcome in breast carcinoma: a comparison of different counting methods in patients with different lymph node status.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 1991, **4:**456-460.
294. Groenendijk RP, Bult P, Noppen CM, Boetes C, Ruers TJ, Wobbes T: **Mitotic activity index in interval breast cancers.** *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 2003, **29:**29-31.
295. Laroye GJ, Panzarella T: **Two techniques for measuring invasion in solid tumors. Evaluated in a retrospective study of 73 cases of breast carcinoma with 10-year follow-up.** *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* 1994, **102:**103-111.
296. van Diest PJ, Matze-Cok E, Baak JP: **Prognostic value of proliferative activity in lymph node metastases of patients with breast cancer.** *Journal of clinical pathology* 1991, **44:**416-418.
297. Takashima T, Onoda N, Ishikawa T, Ogawa Y, Kato Y, Fujimoto Y, Sowa M, Hirakawa K: **Prognostic value of combined analysis of estrogen receptor**

status and cellular proliferative activity in breast cancer patients with extensive lymph node metastases. *Oncology reports* 2002, **9:**589-594.

298. Hasebe T, Iwasaki M, Akashi-Tanaka S, Hojo T, Shibata T, Sasajima Y, Kinoshita T, Tsuda H: **Prognostic significance of mitotic figures in metastatic mammary ductal carcinoma to the lymph nodes.** *Human pathology* 2011, **42:**1823-1832.
299. Penault-Llorca F, Abrial C, Raoelfils I, Chollet P, Cayre A, Mouret-Reynier MA, Thivat E, Mishellany F, Gimbergues P, Durando X: **Changes and predictive and prognostic value of the mitotic index, Ki-67, cyclin D1, and cyclo-oxygenase-2 in 710 operable breast cancer patients treated with neoadjuvant chemotherapy.** *The oncologist* 2008, **13:**1235-1245.
300. Vincent-Salomon A, Rousseau A, Jouve M, Beuzeboc P, Sigal-Zafrani B, Freneaux P, Rosty C, Nos C, Campana F, Klijanienko J, Al Ghuzlan A, Sastre-Garau X, Breast Cancer Study G: **Proliferation markers predictive of the pathological response and disease outcome of patients with breast carcinomas treated by anthracycline-based preoperative chemotherapy.** *European journal of cancer* 2004, **40:**1502-1508.
301. Amat S, Penault-Llorca F, Cure H, Le Bouedec G, Achard JL, Van Praagh I, Feillel V, Mouret-Reynier MA, Dauplat J, Chollet P: **Scarff-Bloom-Richardson (SBR) grading: a pleiotropic marker of chemosensitivity in invasive ductal breast carcinomas treated by neoadjuvant chemotherapy.** *International journal of oncology* 2002, **20:**791-796.
302. Gerdes J, Li L, Schlueter C, Duchrow M, Wohlenberg C, Gerlach C, Stahmer I, Kloth S, Brandt E, Flad HD: **Immunobiochemical and molecular biologic characterization of the cell proliferation-associated nuclear antigen that is defined by monoclonal antibody Ki-67.** *The American journal of pathology* 1991, **138:**867-873.
303. Gerdes J, Schwab U, Lemke H, Stein H: **Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation.** *International journal of cancer Journal international du cancer* 1983, **31:**13-20.
304. Urruticoechea A, Smith IE, Dowsett M: **Proliferation marker Ki-67 in early breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005, **23:**7212-7220.
305. Lopez F, Belloc F, Lacombe F, Dumain P, Reiffers J, Bernard P, Boisseau MR: **Modalities of synthesis of Ki67 antigen during the stimulation of lymphocytes.** *Cytometry* 1991, **12:**42-49.
306. Gerdes J, Lemke H, Baisch H, Wacker HH, Schwab U, Stein H: **Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67.** *Journal of immunology* 1984, **133:**1710-1715.
307. Bruno S, Darzynkiewicz Z: **Cell cycle dependent expression and stability of the nuclear protein detected by Ki-67 antibody in HL-60 cells.** *Cell proliferation* 1992, **25:**31-40.

308. du Manoir S, Guillaud P, Camus E, Seigneurin D, Brugal G: **Ki-67 labeling in postmitotic cells defines different Ki-67 pathways within the 2c compartment.** *Cytometry* 1991, **12:**455-463.
309. Starborg M, Gell K, Brundell E, Hoog C: **The murine Ki-67 cell proliferation antigen accumulates in the nucleolar and heterochromatic regions of interphase cells and at the periphery of the mitotic chromosomes in a process essential for cell cycle progression.** *Journal of cell science* 1996, **109 (Pt 1):**143-153.
310. Booth DG, Takagi M, Sanchez-Pulido L, Petfalski E, Vargiu G, Samejima K, Imamoto N, Ponting CP, Tollervey D, Earnshaw WC, Vagnarelli P: **Ki-67 is a PP1-interacting protein that organises the mitotic chromosome periphery.** *eLife* 2014, **3:**e01641.
311. Gonzalez MA, Tachibana KE, Laskey RA, Coleman N: **Control of DNA replication and its potential clinical exploitation.** *Nature reviews Cancer* 2005, **5:**135-141.
312. Mehrotra P, Gonzalez MA, Johnson SJ, Coleman N, Wilson JA, Davies BR, Lennard TW: **Mcm-2 and Ki-67 have limited potential in preoperative diagnosis of thyroid malignancy.** *The Laryngoscope* 2006, **116:**1434-1438.
313. Clahsen PC, van de Velde CJ, Duval C, Pallud C, Mandard AM, Delobelle-Deroide A, van den Broek L, van de Vijver MJ: **The utility of mitotic index, oestrogen receptor and Ki-67 measurements in the creation of novel prognostic indices for node-negative breast cancer.** *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 1999, **25:**356-363.
314. de Azambuja E, Cardoso F, de Castro G, Jr., Colozza M, Mano MS, Durbecq V, Sotiriou C, Larsimont D, Piccart-Gebhart MJ, Paesmans M: **Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients.** *British journal of cancer* 2007, **96:**1504-1513.
315. Mirza AN, Mirza NQ, Vlastos G, Singletary SE: **Prognostic factors in node-negative breast cancer: a review of studies with sample size more than 200 and follow-up more than 5 years.** *Annals of surgery* 2002, **235:**10-26.
316. Viale G, Giobbie-Hurder A, Regan MM, Coates AS, Mastropasqua MG, Dell'Orto P, Maiorano E, MacGrogan G, Braye SG, Ohlschlegel C, Neven P, Orosz Z, Olszewski WP, Knox F, Thurlimann B, Price KN, Castiglione-Gertsch M, Gelber RD, Gusterson BA, Goldhirsch A, Breast International Group T: **Prognostic and predictive value of centrally reviewed Ki-67 labeling index in postmenopausal women with endocrine-responsive breast cancer: results from Breast International Group Trial 1-98 comparing adjuvant tamoxifen with letrozole.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2008, **26:**5569-5575.
317. Kilickap S, Kaya Y, Yucel B, Tuncer E, Babacan NA, Elagoz S: **Higher Ki67 expression is associates with unfavorable prognostic factors and shorter survival in breast cancer.** *Asian Pacific journal of cancer prevention : APJCP* 2014, **15:**1381-1385.

318. Lau R, Grimson R, Sansome C, Tornos C, Moll UM: **Low levels of cell cycle inhibitor p27kip1 combined with high levels of Ki-67 predict shortened disease-free survival in T1 and T2 invasive breast carcinomas.** *International journal of oncology* 2001, **18:**17-23.
319. Han JS, Cao D, Molberg KH, Sarode VR, Rao R, Sutton LM, Peng Y: **Hormone receptor status rather than HER2 status is significantly associated with increased Ki-67 and p53 expression in triple-negative breast carcinomas, and high expression of Ki-67 but not p53 is significantly associated with axillary nodal metastasis in triple-negative and high-grade non-triple-negative breast carcinomas.** *American journal of clinical pathology* 2011, **135:**230-237.
320. Cabibi D, Mustacchio V, Martorana A, Tripodo C, Campione M, Calascibetta A, Sanguedolce R, Aragona F: **Lymph node metastases displaying lower Ki-67 immunostaining activity than the primary breast cancer.** *Anticancer research* 2006, **26:**4357-4360.
321. Li XR, Liu M, Zhang YJ, Wang JD, Zheng YQ, Li J, Ma B, Song X: **CK5/6, EGFR, Ki-67, cyclin D1, and nm23-H1 protein expressions as predictors of pathological complete response to neoadjuvant chemotherapy in triple-negative breast cancer patients.** *Medical oncology* 2011, **28 Suppl 1:**S129-134.
322. Burcombe RJ, Makris A, Richman PI, Daley FM, Noble S, Pittam M, Wright D, Allen SA, Dove J, Wilson GD: **Evaluation of ER, PgR, HER-2 and Ki-67 as predictors of response to neoadjuvant anthracycline chemotherapy for operable breast cancer.** *British journal of cancer* 2005, **92:**147-155.
323. Assersohn L, Salter J, Powles TJ, A'Hern R, Makris A, Gregory RK, Chang J, Dowsett M: **Studies of the potential utility of Ki67 as a predictive molecular marker of clinical response in primary breast cancer.** *Breast cancer research and treatment* 2003, **82:**113-123.
324. Guarneri V, Piacentini F, Ficarra G, Frassoldati A, D'Amico R, Giovannelli S, Maiorana A, Jovic G, Conte P: **A prognostic model based on nodal status and Ki-67 predicts the risk of recurrence and death in breast cancer patients with residual disease after preoperative chemotherapy.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2009, **20:**1193-1198.
325. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, Griffith C, Boeddinghaus I, Salter J, Detre S, Hills M, Ashley S, Francis S, Walsh G, A'Hern R: **Proliferation and apoptosis as markers of benefit in neoadjuvant endocrine therapy of breast cancer.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2006, **12:**1024s-1030s.
326. Dowsett M, Smith IE, Ebbs SR, Dixon JM, Skene A, A'Hern R, Salter J, Detre S, Hills M, Walsh G, Group IT: **Prognostic value of Ki67 expression after short-term presurgical endocrine therapy for primary breast cancer.** *Journal of the National Cancer Institute* 2007, **99:**167-170.
327. Colozza M, Sidoni A, Piccart-Gebhart M: **Value of Ki67 in breast cancer: the debate is still open.** *The Lancet Oncology* 2010, **11:**414-415.

328. Aleskandarany MA, Rakha EA, Macmillan RD, Powe DG, Ellis IO, Green AR: **MIB1/Ki-67 labelling index can classify grade 2 breast cancer into two clinically distinct subgroups.** *Breast cancer research and treatment* 2011, **127:**591-599.
329. Bevilacqua P, Verderio P, Barbareschi M, Bonoldi E, Boracchi P, Dalla Palma P, Gasparini G: **Lack of prognostic significance of the monoclonal antibody Ki-S1, a novel marker of proliferative activity, in node-negative breast carcinoma.** *Breast cancer research and treatment* 1996, **37:**123-133.
330. Allred DC, Carlson RW, Berry DA, Burstein HJ, Edge SB, Goldstein LJ, Gown A, Hammond ME, Iglehart JD, Moench S, Pierce LJ, Ravdin P, Schnitt SJ, Wolff AC: **NCCN Task Force Report: Estrogen Receptor and Progesterone Receptor Testing in Breast Cancer by Immunohistochemistry.** *Journal of the National Comprehensive Cancer Network : JNCCN* 2009, **7 Suppl 6:**S1-S21; quiz S22-23.
331. Bell SP, Dutta A: **DNA replication in eukaryotic cells.** *Annual review of biochemistry* 2002, **71:**333-374.
332. Labib K, Diffley JF: **Is the MCM2-7 complex the eukaryotic DNA replication fork helicase?** *Current opinion in genetics & development* 2001, **11:**64-70.
333. Chong JP, Mahbubani HM, Khoo CY, Blow JJ: **Purification of an MCM-containing complex as a component of the DNA replication licensing system.** *Nature* 1995, **375:**418-421.
334. Laskey R: **The Croonian Lecture 2001 hunting the antisocial cancer cell: MCM proteins and their exploitation.** *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 2005, **360:**1119-1132.
335. Liang DT, Hodson JA, Forsburg SL: **Reduced dosage of a single fission yeast MCM protein causes genetic instability and S phase delay.** *Journal of cell science* 1999, **112 (Pt 4):**559-567.
336. Bailis JM, Forsburg SL: **MCM proteins: DNA damage, mutagenesis and repair.** *Current opinion in genetics & development* 2004, **14:**17-21.
337. Giaginis C, Georgiadou M, Dimakopoulou K, Tsourouflis G, Gatzidou E, Kouraklis G, Theocharis S: **Clinical significance of MCM-2 and MCM-5 expression in colon cancer: association with clinicopathological parameters and tumor proliferative capacity.** *Digestive diseases and sciences* 2009, **54:**282-291.
338. Toubaji A, Sutcliffe S, Chaux A, Lecksell K, Hicks J, De Marzo AM, Platz EA, Netto GJ: **Immunohistochemical expression of minichromosome maintenance complex protein 2 predicts biochemical recurrence in prostate cancer: a tissue microarray and digital imaging analysis-based study of 428 cases.** *Human pathology* 2012, **43:**1852-1865.
339. Werynska B, Pula B, Muszczynska-Bernhard B, Piotrowska A, Jethon A, Podhorska-Okolow M, Dziegiel P, Jankowska R: **Correlation between expression of metallothionein and expression of Ki-67 and MCM-2**

proliferation markers in non-small cell lung cancer. *Anticancer research* 2011, **31:**2833-2839.

340. Hou Y, Wang HQ, Fu K, Zhang HL, Qian ZZ, Qiu LH, Li W, Zhou SY, Li LF, Hao XS: **[Expression of Cdc7 and mcm2 as a marker for proliferation and prognosis in diffuse large B cell lymphoma].** *Zhonghua zhong liu za zhi [Chinese journal of oncology]* 2011, **33:**911-915.
341. Ramnath N, Hernandez FJ, Tan DF, Huberman JA, Natarajan N, Beck AF, Hyland A, Todorov IT, Brooks JS, Bepler G: **MCM2 is an independent predictor of survival in patients with non-small-cell lung cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2001, **19:**4259-4266.
342. Marshall AE, Rushbrook SM, Vowler SL, Palmer CR, Davies RJ, Gibbs P, Davies SE, Coleman N, Alexander GJ: **Tumor recurrence following liver transplantation for hepatocellular carcinoma: role of tumor proliferation status.** *Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society* 2010, **16:**279-288.
343. Cunha IW, Carvalho KC, Martins WK, Marques SM, Muto NH, Falzoni R, Rocha RM, Aguiar S, Simoes AC, Fahham L, Neves EJ, Soares FA, Reis LF: **Identification of genes associated with local aggressiveness and metastatic behavior in soft tissue tumors.** *Translational oncology* 2010, **3:**23-32.
344. Gonzalez MA, Pinder SE, Callagy G, Vowler SL, Morris LS, Bird K, Bell JA, Laskey RA, Coleman N: **Minichromosome maintenance protein 2 is a strong independent prognostic marker in breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2003, **21:**4306-4313.
345. Mukherjee G, Freeman A, Moore R, Kumaraswamy, Devi KU, Morris LS, Coleman N, Dilworth S, Prabhakaran PS, Stanley MA: **Biologic factors and response to radiotherapy in carcinoma of the cervix.** *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* 2001, **11:**187-193.
346. Loddo M, Kingsbury SR, Rashid M, Proctor I, Holt C, Young J, El-Sheikh S, Falzon M, Eward KL, Prevost T, Sainsbury R, Stoeber K, Williams GH: **Cell-cycle-phase progression analysis identifies unique phenotypes of major prognostic and predictive significance in breast cancer.** *British journal of cancer* 2009, **100:**959-970.
347. Scholzen T, Gerdes J: **The Ki-67 protein: from the known and the unknown.** *Journal of cellular physiology* 2000, **182:**311-322.
348. Cho Mar K, Eimoto T, Nagaya S, Tateyama H: **Cell proliferation marker MCM2, but not Ki67, is helpful for distinguishing between minimally invasive follicular carcinoma and follicular adenoma of the thyroid.** *Histopathology* 2006, **48:**801-807.
349. Guzinska-Ustymowicz K, Pryczynicz A, Kemona A, Czyzewska J: **Correlation between proliferation markers: PCNA, Ki-67, MCM-2 and antiapoptotic**

protein Bcl-2 in colorectal cancer. *Anticancer research* 2009, **29:**3049-3052.

350. Meng MV, Grossfeld GD, Williams GH, Dilworth S, Stoeber K, Mulley TW, Weinberg V, Carroll PR, Tlsty TD: **Minichromosome maintenance protein 2 expression in prostate: characterization and association with outcome after therapy for cancer.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2001, **7:**2712-2718.
351. Wojnar A, Kobierzycki C, Krolicka A, Pula B, Podhorska-Okolow M, Dziegiel P: **Correlation of Ki-67 and MCM-2 proliferative marker expression with grade of histological malignancy (G) in ductal breast cancers.** *Folia histochemica et cytobiologica / Polish Academy of Sciences, Polish Histochemical and Cytochemical Society* 2010, **48:**442-446.
352. Lim LH, Pervaiz S: **Annexin 1: the new face of an old molecule.** *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2007, **21:**968-975.
353. Lindgren CM, Nilsson A, Orho-Melander M, Almgren P, Groop LC: **Characterization of the annexin I gene and evaluation of its role in type 2 diabetes.** *Diabetes* 2001, **50:**2402-2405.
354. de Coupade C, Solito E, Levine JD: **Dexamethasone enhances interaction of endogenous annexin 1 with L-selectin and triggers shedding of L-selectin in the monocytic cell line U-937.** *British journal of pharmacology* 2003, **140:**133-145.
355. Rhee HJ, Kim GY, Huh JW, Kim SW, Na DS: **Annexin I is a stress protein induced by heat, oxidative stress and a sulfhydryl-reactive agent.** *European journal of biochemistry / FEBS* 2000, **267:**3220-3225.
356. Gerke V, Moss SE: **Annexins: from structure to function.** *Physiological reviews* 2002, **82:**331-371.
357. Zhu F, Xu C, Jiang Z, Jin M, Wang L, Zeng S, Teng L, Cao J: **Nuclear localization of annexin A1 correlates with advanced disease and peritoneal dissemination in patients with gastric carcinoma.** *Anatomical record* 2010, **293:**1310-1314.
358. Kalinec F, Webster P, Maricle A, Guerrero D, Chakravarti DN, Chakravarti B, Gellibolian R, Kalinec G: **Glucocorticoid-stimulated, transcription-independent release of annexin A1 by cochlear Hensen cells.** *British journal of pharmacology* 2009, **158:**1820-1834.
359. Ang EZ, Nguyen HT, Sim HL, Putti TC, Lim LH: **Annexin-1 regulates growth arrest induced by high levels of estrogen in MCF-7 breast cancer cells.** *Molecular cancer research : MCR* 2009, **7:**266-274.
360. Hsiang CH, Tunoda T, Whang YE, Tyson DR, Ornstein DK: **The impact of altered annexin I protein levels on apoptosis and signal transduction pathways in prostate cancer cells.** *The Prostate* 2006, **66:**1413-1424.
361. Bizzarro V, Fontanella B, Franceschelli S, Pirozzi M, Christian H, Parente L, Petrella A: **Role of Annexin A1 in mouse myoblast cell differentiation.** *Journal of cellular physiology* 2010, **224:**757-765.

362. Babbin BA, Lee WY, Parkos CA, Winfree LM, Akyildiz A, Perretti M, Nusrat A: **Annexin I regulates SKCO-15 cell invasion by signaling through formyl peptide receptors.** *The Journal of biological chemistry* 2006, **281:**19588-19599.
363. Swa HL, Blackstock WP, Lim LH, Gunaratne J: **Quantitative proteomics profiling of murine mammary gland cells unravels impact of annexin-1 on DNA damage response, cell adhesion, and migration.** *Molecular & cellular proteomics : MCP* 2012, **11:**381-393.
364. Bai XF, Ni XG, Zhao P, Liu SM, Wang HX, Guo B, Zhou LP, Liu F, Zhang JS, Wang K, Xie YQ, Shao YF, Zhao XH: **Overexpression of annexin 1 in pancreatic cancer and its clinical significance.** *World journal of gastroenterology : WJG* 2004, **10:**1466-1470.
365. de Coupade C, Gillet R, Bennoun M, Briand P, Russo-Marie F, Solito E: **Annexin 1 expression and phosphorylation are upregulated during liver regeneration and transformation in antithrombin III SV40 T large antigen transgenic mice.** *Hepatology* 2000, **31:**371-380.
366. Sato Y, Kumamoto K, Saito K, Okayama H, Hayase S, Kofunato Y, Miyamoto K, Nakamura I, Ohki S, Koyama Y, Takenoshita S: **Up-regulated Annexin A1 expression in gastrointestinal cancer is associated with cancer invasion and lymph node metastasis.** *Experimental and therapeutic medicine* 2011, **2:**239-243.
367. Xin W, Rhodes DR, Ingold C, Chinnaiyan AM, Rubin MA: **Dysregulation of the annexin family protein family is associated with prostate cancer progression.** *The American journal of pathology* 2003, **162:**255-261.
368. Petrella A, Festa M, Ercolino SF, Zerilli M, Stassi G, Solito E, Parente L: **Annexin-1 downregulation in thyroid cancer correlates to the degree of tumor differentiation.** *Cancer biology & therapy* 2006, **5:**643-647.
369. Pencil SD, Toth M: **Elevated levels of annexin I protein in vitro and in vivo in rat and human mammary adenocarcinoma.** *Clinical & experimental metastasis* 1998, **16:**113-121.
370. Ahn SH, Sawada H, Ro JY, Nicolson GL: **Differential expression of annexin I in human mammary ductal epithelial cells in normal and benign and malignant breast tissues.** *Clinical & experimental metastasis* 1997, **15:**151-156.
371. Wang LP, Bi J, Yao C, Xu XD, Li XX, Wang SM, Li ZL, Zhang DY, Wang M, Chang GQ: **Annexin A1 expression and its prognostic significance in human breast cancer.** *Neoplasma* 2010, **57:**253-259.
372. Shen D, Nooraie F, Elshimali Y, Lonsberry V, He J, Bose S, Chia D, Seligson D, Chang HR, Goodglick L: **Decreased expression of annexin A1 is correlated with breast cancer development and progression as determined by a tissue microarray analysis.** *Human pathology* 2006, **37:**1583-1591.
373. Cao Y, Li Y, Edelweiss M, Arun B, Rosen D, Resetkova E, Wu Y, Liu J, Sahin A, Albarracin CT: **Loss of annexin A1 expression in breast cancer progression.** *Applied immunohistochemistry & molecular morphology : AIMM*

/ official publication of the Society for Applied Immunohistochemistry 2008, **16:**530-534.

374. Okano M, Kumamoto K, Saito M, Onozawa H, Saito K, Abe N, Ohtake T, Takenoshita S: **Upregulated Annexin A1 promotes cellular invasion in triple-negative breast cancer.** *Oncology reports* 2015, **33:**1064-1070.
375. de Graauw M, van Miltenburg MH, Schmidt MK, Pont C, Lalai R, Kartopawiro J, Pardali E, Le Devedec SE, Smit VT, van der Wal A, Van't Veer LJ, Cleton-Jansen AM, ten Dijke P, van de Water B: **Annexin A1 regulates TGF-beta signaling and promotes metastasis formation of basal-like breast cancer cells.** *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107:**6340-6345.
376. Maschler S, Gebeshuber CA, Wiedemann EM, Alacakaptan M, Schreiber M, Custic I, Beug H: **Annexin A1 attenuates EMT and metastatic potential in breast cancer.** *EMBO molecular medicine* 2010, **2:**401-414.
377. Rong B, Zhao C, Liu H, Ming Z, Cai X, Gao W, Yang S: **Elevated serum annexin A1 as potential diagnostic marker for lung cancer: a retrospective case-control study.** *American journal of translational research* 2014, **6:**558-569.
378. Yu S, Meng Q, Hu H, Zhang M: **Correlation of ANXA1 expression with drug resistance and relapse in bladder cancer.** *International journal of clinical and experimental pathology* 2014, **7:**5538-5548.
379. Yom CK, Han W, Kim SW, Kim HS, Shin HC, Chang JN, Koo M, Noh DY, Moon BI: **Clinical significance of annexin A1 expression in breast cancer.** *Journal of breast cancer* 2011, **14:**262-268.
380. Sano H, Wada S, Eguchi H, Osaki A, Saeki T, Nishiyama M: **Quantitative prediction of tumor response to neoadjuvant chemotherapy in breast cancer: novel marker genes and prediction model using the expression levels.** *Breast cancer* 2012, **19:**37-45.
381. Klein T, Bischoff R: **Physiology and pathophysiology of matrix metalloproteases.** *Amino acids* 2011, **41:**271-290.
382. Stuelten CH, DaCosta Byfield S, Arany PR, Karpova TS, Stetler-Stevenson WG, Roberts AB: **Breast cancer cells induce stromal fibroblasts to express MMP-9 via secretion of TNF-alpha and TGF-beta.** *Journal of cell science* 2005, **118:**2143-2153.
383. Munaut C, Salonurmi T, Kontusaari S, Reponen P, Morita T, Foidart JM, Tryggvason K: **Murine matrix metalloproteinase 9 gene. 5'-upstream region contains cis-acting elements for expression in osteoclasts and migrating keratinocytes in transgenic mice.** *The Journal of biological chemistry* 1999, **274:**5588-5596.
384. Roomi MW, Monterrey JC, Kalinovsky T, Rath M, Niedzwiecki A: **Distinct patterns of matrix metalloproteinase-2 and -9 expression in normal human cell lines.** *Oncology reports* 2009, **21:**821-826.
385. Morelli C, Campioni K, Parolin C, Palu G, Tognon M: **Activity of the matrix metalloproteinase-9 promoter in human normal and tumor cells.** *Journal of cellular physiology* 2004, **199:**126-133.

386. Saito K, Takeha S, Shiba K, Matsuno S, Sorsa T, Nagura H, Ohtani H: **Clinicopathologic significance of urokinase receptor- and MMP-9-positive stromal cells in human colorectal cancer: functional multiplicity of matrix degradation on hematogenous metastasis.** *International journal of cancer Journal international du cancer* 2000, **86:**24-29.
387. Bendeck MP: **Macrophage matrix metalloproteinase-9 regulates angiogenesis in ischemic muscle.** *Circulation research* 2004, **94:**138-139.
388. Yabluchanskiy A, Ma Y, Iyer RP, Hall ME, Lindsey ML: **Matrix metalloproteinase-9: Many shades of function in cardiovascular disease.** *Physiology* 2013, **28:**391-403.
389. Gorden DL, Fingleton B, Crawford HC, Jansen DE, Lepage M, Matrisian LM: **Resident stromal cell-derived MMP-9 promotes the growth of colorectal metastases in the liver microenvironment.** *International journal of cancer Journal international du cancer* 2007, **121:**495-500.
390. Coussens LM, Tinkle CL, Hanahan D, Werb Z: **MMP-9 supplied by bone marrow-derived cells contributes to skin carcinogenesis.** *Cell* 2000, **103:**481-490.
391. Song J, Su H, Zhou YY, Guo LL: **Prognostic value of matrix metalloproteinase 9 expression in breast cancer patients: a meta-analysis.** *Asian Pacific journal of cancer prevention : APJCP* 2013, **14:**1615-1621.
392. Rundhaug JE: **Matrix metalloproteinases, angiogenesis, and cancer: commentary re: A. C. Lockhart et al., Reduction of wound angiogenesis in patients treated with BMS-275291, a broad spectrum matrix metalloproteinase inhibitor. Clin. Cancer Res., 9: 00-00, 2003.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2003, **9:**551-554.
393. Duffy MJ, Maguire TM, Hill A, McDermott E, O'Higgins N: **Metalloproteinases: role in breast carcinogenesis, invasion and metastasis.** *Breast cancer research : BCR* 2000, **2:**252-257.
394. Deryugina EI, Quigley JP: **Matrix metalloproteinases and tumor metastasis.** *Cancer metastasis reviews* 2006, **25:**9-34.
395. Yu W, Liu J, Xiong X, Ai Y, Wang H: **Expression of MMP9 and CD147 in invasive squamous cell carcinoma of the uterine cervix and their implication.** *Pathology, research and practice* 2009, **205:**709-715.
396. Zeng ZS, Huang Y, Cohen AM, Guillem JG: **Prediction of colorectal cancer relapse and survival via tissue RNA levels of matrix metalloproteinase-9.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 1996, **14:**3133-3140.
397. Sillanpaa S, Anttila M, Voutilainen K, Ropponen K, Turpeenniemi-Hujanen T, Puistola U, Tammi R, Tammi M, Sironen R, Saarikoski S, Kosma VM: **Prognostic significance of matrix metalloproteinase-9 (MMP-9) in epithelial ovarian cancer.** *Gynecologic oncology* 2007, **104:**296-303.

398. McGowan PM, Duffy MJ: **Matrix metalloproteinase expression and outcome in patients with breast cancer: analysis of a published database.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2008, **19:**1566-1572.
399. Roy R, Yang J, Moses MA: **Matrix metalloproteinases as novel biomarkers and potential therapeutic targets in human cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009, **27:**5287-5297.
400. Fisher LW, Jain A, Tayback M, Fedarko NS: **Small integrin binding ligand N-linked glycoprotein gene family expression in different cancers.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004, **10:**8501-8511.
401. Zhao M, Hu HG, Huang J, Zou Q, Wang J, Liu MQ, Zhao Y, Li GZ, Xue S, Wu ZS: **Expression and correlation of Twist and gelatinases in breast cancer.** *Experimental and therapeutic medicine* 2013, **6:**97-100.
402. Daniele A, Zito AF, Giannelli G, Divella R, Asselti M, Mazzocca A, Paradiso A, Quaranta M: **Expression of metalloproteinases MMP-2 and MMP-9 in sentinel lymph node and serum of patients with metastatic and non-metastatic breast cancer.** *Anticancer research* 2010, **30:**3521-3527.
403. Zhao S, Ma W, Zhang M, Tang D, Shi Q, Xu S, Zhang X, Liu Y, Song Y, Liu L, Zhang Q: **High expression of CD147 and MMP-9 is correlated with poor prognosis of triple-negative breast cancer (TNBC) patients.** *Medical oncology* 2013, **30:**335.
404. La Rocca G, Pucci-Minafra I, Marrazzo A, Taormina P, Minafra S: **Zymographic detection and clinical correlations of MMP-2 and MMP-9 in breast cancer sera.** *British journal of cancer* 2004, **90:**1414-1421.
405. Wu ZS, Wu Q, Yang JH, Wang HQ, Ding XD, Yang F, Xu XC: **Prognostic significance of MMP-9 and TIMP-1 serum and tissue expression in breast cancer.** *International journal of cancer Journal international du cancer* 2008, **122:**2050-2056.
406. Li HC, Cao DC, Liu Y, Hou YF, Wu J, Lu JS, Di GH, Liu G, Li FM, Ou ZL, Jie C, Shen ZZ, Shao ZM: **Prognostic value of matrix metalloproteinases (MMP-2 and MMP-9) in patients with lymph node-negative breast carcinoma.** *Breast cancer research and treatment* 2004, **88:**75-85.
407. Nolen BM, Marks JR, Ta'san S, Rand A, Luong TM, Wang Y, Blackwell K, Lokshin AE: **Serum biomarker profiles and response to neoadjuvant chemotherapy for locally advanced breast cancer.** *Breast cancer research : BCR* 2008, **10:**R45.
408. AmericanCancerSociety: **American cancer society, Breast Cancer Detailed Guide.** *Atlanta, Ga: American Cancer Society* 2013.
409. He J, Whelan SA, Lu M, Shen D, Chung DU, Saxton RE, Faull KF, Whitelegge JP, Chang HR: **Proteomic-based biosignatures in breast cancer classification and prediction of therapeutic response.** *International journal of proteomics* 2011, **2011:**896476.

410. Gabrovska PS, Robert Anthony; Haupt, Larisa; Griffiths, Lyn: **Gene Expression Profiling in Human Breast Cancer – Toward Personalised Therapeutics?** *The Open Breast Cancer Journal* 2010, **Vol. 2:**pp. 46-59.
411. Garcia Pedrero JM, Fernandez MP, Morgan RO, Herrero Zapatero A, Gonzalez MV, Suarez Nieto C, Rodrigo JP: **Annexin A1 down-regulation in head and neck cancer is associated with epithelial differentiation status.** *The American journal of pathology* 2004, **164:**73-79.
412. Hu N, Flaig MJ, Su H, Shou JZ, Roth MJ, Li WJ, Wang C, Goldstein AM, Li G, Emmert-Buck MR, Taylor PR: **Comprehensive characterization of annexin I alterations in esophageal squamous cell carcinoma.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004, **10:**6013-6022.
413. Sinha P, Hutter G, Kottgen E, Dietel M, Schadendorf D, Lage H: **Increased expression of annexin I and thioredoxin detected by two-dimensional gel electrophoresis of drug resistant human stomach cancer cells.** *Journal of biochemical and biophysical methods* 1998, **37:**105-116.
414. Yi M, Schnitzer JE: **Impaired tumor growth, metastasis, angiogenesis and wound healing in annexin A1-null mice.** *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106:**17886-17891.
415. Jezequel P, Frenel JS, Campion L, Guerin-Charbonnel C, Gouraud W, Ricolleau G, Campone M: **bc-GenExMiner 3.0: new mining module computes breast cancer gene expression correlation analyses.** *Database : the journal of biological databases and curation* 2013, **2013:**bas060.
416. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW: **A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes.** *Cancer cell* 2006, **10:**515-527.
417. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET, Bergh J: **An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102:**13550-13555.
418. Tavassoli FA DP: **World Health Organization classification of tumours. In Pathology and Genetics Tumours of the Breast and Female Genital Organs.** *Lyon: IARC Press;* 2003**:**19-23.
419. Hammond ME, Hayes DF, Wolff AC, Mangu PB, Temin S: **American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer.** *Journal of oncology practice / American Society of Clinical Oncology* 2010, **6:**195-197.
420. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R,

Vance GH, van de Vijver M, Wheeler TM, Hayes DF, American Society of Clinical O, College of American P: **American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007, **25:**118-145.

421. Caldarella A, Puliti D, Crocetti E, Bianchi S, Vezzosi V, Apicella P, Biancalani M, Giannini A, Urso C, Zolfanelli F, Paci E: **Biological characteristics of interval cancers: a role for biomarkers in the breast cancer screening.** *Journal of cancer research and clinical oncology* 2013, **139:**181-185.
422. Subik K, Lee JF, Baxter L, Strzepek T, Costello D, Crowley P, Xing L, Hung MC, Bonfiglio T, Hicks DG, Tang P: **The Expression Patterns of ER, PR, HER2, CK5/6, EGFR, Ki-67 and AR by Immunohistochemical Analysis in Breast Cancer Cell Lines.** *Breast cancer : basic and clinical research* 2010, **4:**35-41.
423. Jonsson G, Staaf J, Olsson E, Heidenblad M, Vallon-Christersson J, Osoegawa K, de Jong P, Oredsson S, Ringner M, Hoglund M, Borg A: **High-resolution genomic profiles of breast cancer cell lines assessed by tiling BAC array comparative genomic hybridization.** *Genes, chromosomes & cancer* 2007, **46:**543-558.
424. Elshimali Y, Liu P: **Expression Of Annexin A1 in Normal and Malignant Breast Epithelium and Its Clinical Significance.** *Calcium Binding Proteins* 2006, **1:**51-53.
425. Kang H, Ko J, Jang SW: **The role of annexin A1 in expression of matrix metalloproteinase-9 and invasion of breast cancer cells.** *Biochemical and biophysical research communications* 2012, **423:**188-194.
426. Creighton CJ: **The molecular profile of luminal B breast cancer.** *Biologics : targets & therapy* 2012, **6:**289-297.
427. Wang CC, Liau JY, Lu YS, Chen JW, Yao YT, Lien HC: **Differential expression of moesin in breast cancers and its implication in epithelial-mesenchymal transition.** *Histopathology* 2012, **61:**78-87.
428. Choi YL, Bocanegra M, Kwon MJ, Shin YK, Nam SJ, Yang JH, Kao J, Godwin AK, Pollack JR: **LYN is a mediator of epithelial-mesenchymal transition and a target of dasatinib in breast cancer.** *Cancer research* 2010, **70:**2296-2306.
429. Bernardo GM, Bebek G, Ginther CL, Sizemore ST, Lozada KL, Miedler JD, Anderson LA, Godwin AK, Abdul-Karim FW, Slamon DJ, Keri RA: **FOXA1 represses the molecular phenotype of basal breast cancer cells.** *Oncogene* 2013, **32:**554-563.
430. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS: **Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4.** *Molecular cell* 2002, **9:**279-289.
431. Nakshatri H, Badve S: **FOXA1 in breast cancer.** *Expert reviews in molecular medicine* 2009, **11:**e8.
432. Robinson JL, Carroll JS: **FoxA1 is a key mediator of hormonal response in breast and prostate cancer.** *Frontiers in endocrinology* 2012, **3:**68.

433. Guo BH, Feng Y, Zhang R, Xu LH, Li MZ, Kung HF, Song LB, Zeng MS: **Bmi-1 promotes invasion and metastasis, and its elevated expression is correlated with an advanced stage of breast cancer.** *Molecular cancer* 2011, **10:**10.
434. Stein T, Price KN, Morris JS, Heath VJ, Ferrier RK, Bell AK, Pringle MA, Villadsen R, Petersen OW, Sauter G, Bryson G, Mallon EA, Gusterson BA: **Annexin A8 is up-regulated during mouse mammary gland involution and predicts poor survival in breast cancer.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2005, **11:**6872-6879.
435. Ibrahim T, Mercatali L, Amadori D: **A new emergency in oncology: Bone metastases in breast cancer patients (Review).** *Oncology letters* 2013, **6:**306-310.
436. Ravnan MPWaSL: **Metastatic breast cancer: A review of current and novel pharmacotherapy.** *Formulary* 2011, **46:**130–146.
437. Gupta GP, Massague J: **Cancer metastasis: building a framework.** *Cell* 2006, **127:**679-695.
438. Nguyen DX, Bos PD, Massague J: **Metastasis: from dissemination to organ-specific colonization.** *Nature reviews Cancer* 2009, **9:**274-284.
439. Hsiao KC, Shih NY, Fang HL, Huang TS, Kuo CC, Chu PY, Hung YM, Chou SW, Yang YY, Chang GC, Liu KJ: **Surface alpha-enolase promotes extracellular matrix degradation and tumor metastasis and represents a new therapeutic target.** *PloS one* 2013, **8:**e69354.
440. Van den Steen PE, Proost P, Brand DD, Kang AH, Van Damme J, Opdenakker G: **Generation of glycosylated remnant epitopes from human collagen type II by gelatinase B.** *Biochemistry* 2004, **43:**10809-10816.
441. Rosenblum G, Van den Steen PE, Cohen SR, Bitler A, Brand DD, Opdenakker G, Sagi I: **Direct visualization of protease action on collagen triple helical structure.** *PloS one* 2010, **5:**e11043.
442. Sand JM, Larsen L, Hogaboam C, Martinez F, Han M, Rossel Larsen M, Nawrocki A, Zheng Q, Karsdal MA, Leeming DJ: **MMP mediated degradation of type IV collagen alpha 1 and alpha 3 chains reflects basement membrane remodeling in experimental and clinical fibrosis--validation of two novel biomarker assays.** *PloS one* 2013, **8:**e84934.
443. Kessenbrock K, Plaks V, Werb Z: **Matrix metalloproteinases: regulators of the tumor microenvironment.** *Cell* 2010, **141:**52-67.
444. Tavassoli FA DP: **World Health Organization classification of tumours. In Pathology and Genetics Tumours of the Breast and Female Genital Organs.** *Lyon: IARC Press* 2003**:**p. 19-23.
445. Jung IH, Jung DE, Park YN, Song SY, Park SW: **Aberrant Hedgehog ligands induce progressive pancreatic fibrosis by paracrine activation of myofibroblasts and ductular cells in transgenic zebrafish.** *PloS one* 2011, **6:**e27941.
446. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu

PB, Paik S, Perez EA, Press MF, Spears PA, Vance GH, Viale G, Hayes DF: **Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Update.** *Archives of pathology & laboratory medicine* 2013.

447. Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, Kwei KA, Hernandez-Boussard T, Wang P, Gazdar AF, Minna JD, Pollack JR: **Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery.** *PloS one* 2009, **4:**e6146.
448. Wilson TR, Fridlyand J, Yan Y, Penuel E, Burton L, Chan E, Peng J, Lin E, Wang Y, Sosman J, Ribas A, Li J, Moffat J, Sutherlin DP, Koeppen H, Merchant M, Neve R, Settleman J: **Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors.** *Nature* 2012, **487:**505-509.
449. Hoeflich KP, O'Brien C, Boyd Z, Cavet G, Guerrero S, Jung K, Januario T, Savage H, Punnoose E, Truong T, Zhou W, Berry L, Murray L, Amler L, Belvin M, Friedman LS, Lackner MR: **In vivo antitumor activity of MEK and phosphatidylinositol 3-kinase inhibitors in basal-like breast cancer models.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2009, **15:**4649-4664.
450. Kousidou OC, Roussidis AE, Theocharis AD, Karamanos NK: **Expression of MMPs and TIMPs genes in human breast cancer epithelial cells depends on cell culture conditions and is associated with their invasive potential.** *Anticancer research* 2004, **24:**4025-4030.
451. Yao J, Xiong S, Klos K, Nguyen N, Grijalva R, Li P, Yu D: **Multiple signaling pathways involved in activation of matrix metalloproteinase-9 (MMP-9) by heregulin-beta1 in human breast cancer cells.** *Oncogene* 2001, **20:**8066-8074.
452. Illemann M, Bird N, Majeed A, Sehested M, Laerum OD, Lund LR, Dano K, Nielsen BS: **MMP-9 is differentially expressed in primary human colorectal adenocarcinomas and their metastases.** *Molecular cancer research : MCR* 2006, **4:**293-302.
453. Egeblad M, Werb Z: **New functions for the matrix metalloproteinases in cancer progression.** *Nature reviews Cancer* 2002, **2:**161-174.
454. Lopez-Otin C, Matrisian LM: **Emerging roles of proteases in tumour suppression.** *Nature reviews Cancer* 2007, **7:**800-808.
455. Cao D, Polyak K, Halushka MK, Nassar H, Kouprina N, Iacobuzio-Donahue C, Wu X, Sukumar S, Hicks J, De Marzo A, Argani P: **Serial analysis of gene expression of lobular carcinoma in situ identifies down regulation of claudin 4 and overexpression of matrix metalloproteinase 9.** *Breast cancer research : BCR* 2008, **10:**R91.
456. St-Pierre Y, Couillard J, Van Themsche C: **Regulation of MMP-9 gene expression for the development of novel molecular targets against cancer and inflammatory diseases.** *Expert opinion on therapeutic targets* 2004, **8:**473-489.

457. Radenkovic S, Konjevic G, Jurisic V, Karadzic K, Nikitovic M, Gopcevic K: **Values of MMP-2 and MMP-9 in Tumor Tissue of Basal-Like Breast Cancer Patients.** *Cell biochemistry and biophysics* 2014, **68:**143-152.
458. Mehner C, Hockla A, Miller E, Ran S, Radisky DC, Radisky ES: **Tumor cell-produced matrix metalloproteinase 9 (MMP-9) drives malignant progression and metastasis of basal-like triple negative breast cancer.** *Oncotarget* 2014, **5:**2736-2749.
459. Labrie M, St-Pierre Y: **Epigenetic regulation of mmp-9 gene expression.** *Cellular and molecular life sciences : CMLS* 2013, **70:**3109-3124.
460. Crowe DL, Brown TN: **Transcriptional inhibition of matrix metalloproteinase 9 (MMP-9) activity by a c-fos/estrogen receptor fusion protein is mediated by the proximal AP-1 site of the MMP-9 promoter and correlates with reduced tumor cell invasion.** *Neoplasia* 1999, **1:**368-372.
461. Orlichenko LS, Radisky DC: **Matrix metalloproteinases stimulate epithelial-mesenchymal transition during tumor development.** *Clinical & experimental metastasis* 2008, **25:**593-600.
462. David JM, Rajasekaran AK: **Dishonorable discharge: the oncogenic roles of cleaved E-cadherin fragments.** *Cancer research* 2012, **72:**2917-2923.
463. Lin CY, Tsai PH, Kandaswami CC, Lee PP, Huang CJ, Hwang JJ, Lee MT: **Matrix metalloproteinase-9 cooperates with transcription factor Snail to induce epithelial-mesenchymal transition.** *Cancer science* 2011, **102:**815-827.
464. Wu QW, Yang QM, Huang YF, She HQ, Liang J, Yang QL, Zhang ZM: **Expression and clinical significance of matrix metalloproteinase-9 in lymphatic invasiveness and metastasis of breast cancer.** *PloS one* 2014, **9:**e97804.
465. Ross JS, Hatzis C, Symmans WF, Pusztai L, Hortobagyi GN: **Commercialized multigene predictors of clinical outcome for breast cancer.** *The oncologist* 2008, **13:**477-493.
466. Viale G: **The current state of breast cancer classification.** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2012, **23 Suppl 10:**x207-210.
467. Sihto H, Lundin J, Lehtimaki T, Sarlomo-Rikala M, Butzow R, Holli K, Sailas L, Kataja V, Lundin M, Turpeenniemi-Hujanen T, Isola J, Heikkila P, Joensuu H: **Molecular subtypes of breast cancers detected in mammography screening and outside of screening.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2008, **14:**4103-4110.
468. Weigelt B, Baehner FL, Reis-Filho JS: **The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade.** *The Journal of pathology* 2010, **220:**263-280.
469. Whitfield ML, George LK, Grant GD, Perou CM: **Common markers of proliferation.** *Nature reviews Cancer* 2006, **6:**99-106.

470. Ganguly A, Shields CL: **Differential gene expression profile of retinoblastoma compared to normal retina.** *Molecular vision* 2010, **16:**1292-1303.
471. Cuzick J, Swanson GP, Fisher G, Brothman AR, Berney DM, Reid JE, Mesher D, Speights VO, Stankiewicz E, Foster CS, Moller H, Scardino P, Warren JD, Park J, Younus A, Flake DD, 2nd, Wagner S, Gutin A, Lanchbury JS, Stone S, Transatlantic Prostate G: **Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study.** *The Lancet Oncology* 2011, **12:**245-255.
472. Andrisani OM, Studach L, Merle P: **Gene signatures in hepatocellular carcinoma (HCC).** *Seminars in cancer biology* 2011, **21:**4-9.
473. Wistuba, II, Behrens C, Lombardi F, Wagner S, Fujimoto J, Raso MG, Spaggiari L, Galetta D, Riley R, Hughes E, Reid J, Sangale Z, Swisher SG, Kalhor N, Moran CA, Gutin A, Lanchbury JS, Barberis M, Kim ES: **Validation of a proliferation-based expression signature as prognostic marker in early stage lung adenocarcinoma.** *Clinical cancer research : an official journal of the American Association for Cancer Research* 2013, **19:**6261-6271.
474. van Diest PJ, van der Wall E, Baak JP: **Prognostic value of proliferation in invasive breast cancer: a review.** *Journal of clinical pathology* 2004, **57:**675-681.
475. Dai H, van't Veer L, Lamb J, He YD, Mao M, Fine BM, Bernards R, van de Vijver M, Deutsch P, Sachs A, Stoughton R, Friend S: **A cell proliferation signature is a marker of extremely poor outcome in a subpopulation of breast cancer patients.** *Cancer research* 2005, **65:**4059-4066.
476. Klein ME, Dabbs DJ, Shuai Y, Brufsky AM, Jankowitz R, Puhalla SL, Bhargava R: **Prediction of the Oncotype DX recurrence score: use of pathology-generated equations derived by linear regression analysis.** *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2013, **26:**658-664.
477. Chibon F: **Cancer gene expression signatures - the rise and fall?** *European journal of cancer* 2013, **49:**2000-2009.
478. Beresford MJ, Wilson GD, Makris A: **Measuring proliferation in breast cancer: practicalities and applications.** *Breast cancer research : BCR* 2006, **8:**216.
479. Colozza M, Azambuja E, Cardoso F, Sotiriou C, Larsimont D, Piccart MJ: **Proliferative markers as prognostic and predictive tools in early breast cancer: where are we now?** *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 2005, **16:**1723-1739.
480. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA: **Ki67 in breast cancer: prognostic and predictive potential.** *The Lancet Oncology* 2010, **11:**174-183.
481. Wharton SB, Hibberd S, Eward KL, Crimmins D, Jellinek DA, Levy D, Stoeber K, Williams GH: **DNA replication licensing and cell cycle kinetics of oligodendroglial tumours.** *British journal of cancer* 2004, **91:**262-269.

482. Stoeber K, Tlsty TD, Happerfield L, Thomas GA, Romanov S, Bobrow L, Williams ED, Williams GH: **DNA replication licensing and human cell proliferation.** *Journal of cell science* 2001, **114:**2027-2041.
483. Yang C, Wen Y, Li H, Zhang D, Zhang N, Shi X, Jiang B, Ma X, Yang P, Tang H, Peng Z, Yang Y: **Overexpression of minichromosome maintenance 2 predicts poor prognosis in patients with gastric cancer.** *Oncology reports* 2012, **27:**135-142.
484. Liu Y, He G, Wang Y, Guan X, Pang X, Zhang B: **MCM-2 is a therapeutic target of Trichostatin A in colon cancer cells.** *Toxicology letters* 2013, **221:**23-30.
485. Abdou AG, Elwahed MG, Serag El-Dien MM, Eldien DS: **Immunohistochemical Expression of MCM2 in Nonmelanoma Epithelial Skin Cancers.** *The American Journal of dermatopathology* 2014.
486. Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, Harte RA, Heitner S, Hickey G, Hinrichs AS, Hubley R, Karolchik D, Learned K, Lee BT, Li CH, Miga KH, Nguyen N, Paten B, Raney BJ, Smit AF, Speir ML, Zweig AS, Haussler D, Kuhn RM, Kent WJ: **The UCSC Genome Browser database: 2015 update.** *Nucleic acids research* 2015, **43:**D670-681.
487. Wang Y, Zhou ZG, Xia QJ, Zhang WY, Li HG, Wang R: **[Expression of minichromosome maintenance protein 2 in colonic adenocarcinoma, adenoma and normal colonic mucosa and its clinical significance].** *Zhonghua wei chang wai ke za zhi = Chinese journal of gastrointestinal surgery* 2008, **11:**465-468.
488. Yousef EM, Tahir MR, St-Pierre Y, Gaboury LA: **MMP-9 expression varies according to molecular subtypes of breast cancer.** *BMC cancer* 2014, **14:**609.
489. Malhotra A, Younes M, Kuna ST, Benca R, Kushida CA, Walsh J, Hanlon A, Staley B, Pack AI, Pien GW: **Performance of an automated polysomnography scoring system versus computer-assisted manual scoring.** *Sleep* 2013, **36:**573-582.
490. Singh G: **Determination of Cutoff Score for a Diagnostic Test.** *the Internet Journal of Laboratory Medicine* 2006, **2**.
491. Wen Zhu NZ, Ning Wang: **Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS® Implementations.** *NESUG* 2010.
492. Yang Z, Tang LH, Klimstra DS: **Effect of tumor heterogeneity on the assessment of Ki67 labeling index in well-differentiated neuroendocrine tumors metastatic to the liver: implications for prognostic stratification.** *The American journal of surgical pathology* 2011, **35:**853-860.
493. Tabata K, Tanaka T, Hayashi T, Hori T, Nunomura S, Yonezawa S, Fukuoka J: **Ki-67 is a strong prognostic marker of non-small cell lung cancer when tissue heterogeneity is considered.** *BMC clinical pathology* 2014, **14:**23.

494. Shetty A, Loddo M, Fanshawe T, Prevost AT, Sainsbury R, Williams GH, Stoeber K: **DNA replication licensing and cell cycle kinetics of normal and neoplastic breast.** *British journal of cancer* 2005, **93:**1295-1300.
495. Alison MR, Hunt T, Forbes SJ: **Minichromosome maintenance (MCM) proteins may be pre-cancer markers.** *Gut* 2002, **50:**290-291.
496. Ali HR, Dawson SJ, Blows FM, Provenzano E, Pharoah PD, Caldas C: **Aurora kinase A outperforms Ki67 as a prognostic marker in ER-positive breast cancer.** *British journal of cancer* 2012, **106:**1798-1806.
497. Keam B, Im SA, Lee KH, Han SW, Oh DY, Kim JH, Lee SH, Han W, Kim DW, Kim TY, Park IA, Noh DY, Heo DS, Bang YJ: **Ki-67 can be used for further classification of triple negative breast cancer into two subtypes with different response and prognosis.** *Breast cancer research : BCR* 2011, **13:**R22.
498. Ribelles N, Perez-Villa L, Jerez JM, Pajares B, Vicioso L, Jimenez B, de Luque V, Franco L, Gallego E, Marquez A, Alvarez M, Sanchez-Munoz A, Perez-Rivas L, Alba E: **Pattern of recurrence of early breast cancer is different according to intrinsic subtype and proliferation index.** *Breast cancer research : BCR* 2013, **15:**R98.
499. Li FY, Wu SG, Zhou J, Sun JY, Lin Q, Lin HX, Guan XX, He ZY: **Prognostic value of Ki-67 in breast cancer patients with positive axillary lymph nodes: a retrospective cohort study.** *PloS one* 2014, **9:**e87264.
500. Ferguson NL, Bell J, Heidel R, Lee S, Vanmeter S, Duncan L, Munsey B, Panella T, Orucevic A: **Prognostic value of breast cancer subtypes, Ki-67 proliferation index, age, and pathologic tumor characteristics on breast cancer survival in Caucasian women.** *The breast journal* 2013, **19:**22-30.
501. Bessarabova M, Kirillov E, Shi W, Bugrim A, Nikolsky Y, Nikolskaya T: **Bimodal gene expression patterns in breast cancer.** *BMC genomics* 2010, **11 Suppl 1:**S8.
502. Blow JJ, Hodgson B: **Replication licensing--defining the proliferative state?** *Trends in cell biology* 2002, **12:**72-78.
503. van Dierendonck JH, Keijzer R, van de Velde CJ, Cornelisse CJ: **Nuclear distribution of the Ki-67 antigen during the cell cycle: comparison with growth fraction in human breast cancer cells.** *Cancer research* 1989, **49:**2999-3006.
504. Heidebrecht HJ, Buck F, Haas K, Wacker HH, Parwaresch R: **Monoclonal antibodies Ki-S3 and Ki-S5 yield new data on the 'Ki-67' proteins.** *Cell proliferation* 1996, **29:**413-425.
505. Luck AA, Evans AJ, Green AR, Rakha EA, Paish C, Ellis IO: **The influence of basal phenotype on the metastatic pattern of breast cancer.** *Clinical oncology* 2008, **20:**40-45.
506. Merdad A, Karim S, Schulten HJ, Dallol A, Buhmeida A, Al-Thubaity F, Gari MA, Chaudhary AG, Abuzenadah AM, Al-Qahtani MH: **Expression of matrix metalloproteinases (MMPs) in primary human breast cancer: MMP-9 as a potential biomarker for cancer invasion and metastasis.** *Anticancer research* 2014, **34:**1355-1366.

507. Arnedos M, Bihan C, Delaloge S, Andre F: **Triple-negative breast cancer: are we making headway at least?** *Therapeutic advances in medical oncology* 2012, **4:**195-210.
508. Su Y, Shrubsole MJ, Ness RM, Cai Q, Kataoka N, Washington K, Zheng W: **Immunohistochemical expressions of Ki-67, cyclin D1, beta-catenin, cyclooxygenase-2, and epidermal growth factor receptor in human colorectal adenoma: a validation study of tissue microarrays.** *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* 2006, **15:**1719-1726.
509. Frampton JP, White JB, Simon AB, Tsuei M, Paczesny S, Takayama S: **Aqueous two-phase system patterning of detection antibody solutions for cross-reaction-free multiplex ELISA.** *Scientific reports* 2014, **4:**4878.
510. Vernon AE, Bakewell SJ, Chodosh LA: **Deciphering the molecular basis of breast cancer metastasis with mouse models.** *Reviews in endocrine & metabolic disorders* 2007, **8:**199-213.
511. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, Gerald WL, Massague J: **Endogenous human microRNAs that suppress breast cancer metastasis.** *Nature* 2008, **451:**147-152.
512. Yao ES, Zhang H, Chen YY, Lee B, Chew K, Moore D, Park C: **Increased beta1 integrin is associated with decreased survival in invasive breast cancer.** *Cancer research* 2007, **67:**659-664.
513. Zani S, Clary BM: **A role for hepatic metastasectomy in stage IV melanoma and breast cancer: reestablishing the surgical modality.** *Oncology* 2011, **25:**1158-1164.
514. DeSantis C, Siegel R, Bandi P, Jemal A: **Breast cancer statistics, 2011.** *CA Cancer J Clin* 2011, **61:**409-418.
515. Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJ, Naghavi M: **Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis.** *Lancet* 2011, **378:**1461-1484.
516. Steeg PS: **Metastasis suppressors alter the signal transduction of cancer cells.** *Nature reviews Cancer* 2003, **3:**55-63.
517. Chaffer CL, Weinberg RA: **A perspective on cancer cell metastasis.** *Science* 2011, **331:**1559-1564.
518. Mego M, Mani SA, Cristofanilli M: **Molecular mechanisms of metastasis in breast cancer--clinical applications.** *Nature reviews Clinical oncology* 2010, **7:**693-701.
519. Onishi T, Hayashi N, Theriault RL, Hortobagyi GN, Ueno NT: **Future directions of bone-targeted therapy for metastatic breast cancer.** *Nature reviews Clinical oncology* 2010, **7:**641-651.
520. Ebos JM, Kerbel RS: **Antiangiogenic therapy: impact on invasion, disease progression, and metastasis.** *Nature reviews Clinical oncology* 2011, **8:**210-221.
521. Kraljevic Pavelic S, Sedic M, Bosnjak H, Spaventi S, Pavelic K: **Metastasis: new perspectives on an old problem.** *Molecular cancer* 2011, **10:**22.

522. Curran S, Murray GI: **Matrix metalloproteinases: molecular aspects of their roles in tumour invasion and metastasis.** *European journal of cancer* 2000, **36:**1621-1630.
523. Meyer T, Hart IR: **Mechanisms of tumour metastasis.** *European journal of cancer* 1998, **34:**214-221.
524. Linder P: **Dead-box proteins: a family affair--active and passive players in RNP-remodeling.** *Nucleic acids research* 2006, **34:**4168-4180.
525. Mouillet JF, Yan X, Ou Q, Jin L, Muglia LJ, Crawford PA, Sadovsky Y: **DEAD-box protein-103 (DP103, Ddx20) is essential for early embryonic development and modulates ovarian morphology and function.** *Endocrinology* 2008, **149:**2168-2175.
526. Fuller-Pace FV: **DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation.** *Nucleic acids research* 2006, **34:**4206-4215.
527. Lee MB, Lebedeva LA, Suzawa M, Wadekar SA, Desclozeaux M, Ingraham HA: **The DEAD-box protein DP103 (Ddx20 or Gemin-3) represses orphan nuclear receptor activity via SUMO modification.** *Molecular and cellular biology* 2005, **25:**1879-1890.
528. Ou Q, Mouillet JF, Yan X, Dorn C, Crawford PA, Sadovsky Y: **The DEAD box protein DP103 is a regulator of steroidogenic factor-1.** *Molecular endocrinology* 2001, **15:**69-79.
529. Yan X, Mouillet JF, Ou Q, Sadovsky Y: **A novel domain within the DEAD-box protein DP103 is essential for transcriptional repression and helicase activity.** *Molecular and cellular biology* 2003, **23:**414-423.
530. Klappacher GW, Lunyak VV, Sykes DB, Sawka-Verhelle D, Sage J, Brard G, Ngo SD, Gangadharan D, Jacks T, Kamps MP, Rose DW, Rosenfeld MG, Glass CK: **An induced Ets repressor complex regulates growth arrest during terminal macrophage differentiation.** *Cell* 2002, **109:**169-180.
531. Charroux B, Pellizzoni L, Perkinson RA, Shevchenko A, Mann M, Dreyfuss G: **Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems.** *The Journal of cell biology* 1999, **147:**1181-1194.
532. Grundhoff AT, Kremmer E, Tureci O, Glieden A, Gindorf C, Atz J, Mueller-Lantzsch N, Schubach WH, Grasser FA: **Characterization of DP103, a novel DEAD box protein that binds to the Epstein-Barr virus nuclear proteins EBNA2 and EBNA3C.** *The Journal of biological chemistry* 1999, **274:**19136-19144.
533. Meister G, Buhler D, Laggerbauer B, Zobawa M, Lottspeich F, Fischer U: **Characterization of a nuclear 20S complex containing the survival of motor neurons (SMN) protein and a specific subset of spliceosomal Sm proteins.** *Human molecular genetics* 2000, **9:**1977-1986.
534. Shpargel KB, Matera AG: **Gemin proteins are required for efficient assembly of Sm-class ribonucleoproteins.** *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102:**17372-17377.

535. Gillian AL, Svaren J: **The Ddx20/DP103 dead box protein represses transcriptional activation by Egr2/Krox-20.** *The Journal of biological chemistry* 2004, **279:**9056-9063.
536. Hayden MS, Ghosh S: **Signaling to NF-kappaB.** *Genes & development* 2004, **18:**2195-2224.
537. Verma IM, Stevenson JK, Schwarz EM, Van Antwerp D, Miyamoto S: **Rel/NF-kappa B/I kappa B family: intimate tales of association and dissociation.** *Genes & development* 1995, **9:**2723-2735.
538. Perkins ND: **Integrating cell-signalling pathways with NF-kappaB and IKK function.** *Nature reviews Molecular cell biology* 2007, **8:**49-62.
539. Bhoj VG, Chen ZJ: **Ubiquitylation in innate and adaptive immunity.** *Nature* 2009, **458:**430-437.
540. Teo H, Ghosh S, Luesch H, Ghosh A, Wong ET, Malik N, Orth A, de Jesus P, Perry AS, Oliver JD, Tran NL, Speiser LJ, Wong M, Saez E, Schultz P, Chanda SK, Verma IM, Tergaonkar V: **Telomere-independent Rap1 is an IKK adaptor and regulates NF-kappaB-dependent gene expression.** *Nature cell biology* 2010, **12:**758-767.
541. Wu ZH, Wong ET, Shi Y, Niu J, Chen Z, Miyamoto S, Tergaonkar V: **ATM- and NEMO-dependent ELKS ubiquitination coordinates TAK1-mediated IKK activation in response to genotoxic stress.** *Molecular cell* 2010, **40:**75-86.
542. Staudt LM: **Oncogenic activation of NF-kappaB.** *Cold Spring Harbor perspectives in biology* 2010, **2:**a000109.
543. Ben-Neriah Y, Karin M: **Inflammation meets cancer, with NF-kappaB as the matchmaker.** *Nature immunology* 2011, **12:**715-723.
544. Ghosh S, Karin M: **Missing pieces in the NF-kappaB puzzle.** *Cell* 2002, **109 Suppl:**S81-96.
545. Bhat-Nakshatri P, Sweeney CJ, Nakshatri H: **Identification of signal transduction pathways involved in constitutive NF-kappaB activation in breast cancer cells.** *Oncogene* 2002, **21:**2066-2078.
546. Nakshatri H, Goulet RJ, Jr.: **NF-kappaB and breast cancer.** *Current problems in cancer* 2002, **26:**282-309.
547. Romieu-Mourez R, Kim DW, Shin SM, Demicco EG, Landesman-Bollag E, Seldin DC, Cardiff RD, Sonenshein GE: **Mouse mammary tumor virus c-rel transgenic mice develop mammary tumors.** *Molecular and cellular biology* 2003, **23:**5738-5754.
548. Perkins ND: **The diverse and complex roles of NF-kappaB subunits in cancer.** *Nature reviews Cancer* 2012, **12:**121-132.
549. Sovak MA, Bellas RE, Kim DW, Zanieski GJ, Rogers AE, Traish AM, Sonenshein GE: **Aberrant nuclear factor-kappaB/Rel expression and the pathogenesis of breast cancer.** *The Journal of clinical investigation* 1997, **100:**2952-2960.
550. Miyamoto S: **Nuclear initiated NF-kappaB signaling: NEMO and ATM take center stage.** *Cell research* 2011, **21:**116-130.

551. Chaturvedi MM, Sung B, Yadav VR, Kannappan R, Aggarwal BB: **NF-kappaB addiction and its role in cancer: 'one size does not fit all'.** *Oncogene* 2011, **30:**1615-1630.
552. Basseres DS, Baldwin AS: **Nuclear factor-kappaB and inhibitor of kappaB kinase pathways in oncogenic initiation and progression.** *Oncogene* 2006, **25:**6817-6830.
553. Kim HJ, Hawke N, Baldwin AS: **NF-kappaB and IKK as therapeutic targets in cancer.** *Cell death and differentiation* 2006, **13:**738-747.
554. Biswas SK, Tergaonkar V: **Myeloid differentiation factor 88-independent Toll-like receptor pathway: Sustaining inflammation or promoting tolerance?** *The international journal of biochemistry & cell biology* 2007, **39:**1582-1592.
555. Rajendran P, Li F, Shanmugam MK, Kannaiyan R, Goh JN, Wong KF, Wang W, Khin E, Tergaonkar V, Kumar AP, Luk JM, Sethi G: **Celastrol suppresses growth and induces apoptosis of human hepatocellular carcinoma through the modulation of STAT3/JAK2 signaling cascade in vitro and in vivo.** *Cancer prevention research* 2012, **5:**631-643.
556. Dey A, Wong E, Kua N, Teo HL, Tergaonkar V, Lane D: **Hexamethylene bisacetamide (HMBA) simultaneously targets AKT and MAPK pathway and represses NF kappaB activity: implications for cancer therapy.** *Cell cycle* 2008, **7:**3759-3767.
557. Irelan JT, Murphy TJ, DeJesus PD, Teo H, Xu D, Gomez-Ferreria MA, Zhou Y, Miraglia LJ, Rines DR, Verma IM, Sharp DJ, Tergaonkar V, Chanda SK: **A role for IkappaB kinase 2 in bipolar spindle assembly.** *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104:**16940-16945.
558. Ang HL, Tergaonkar V: **Notch and NFkappaB signaling pathways: Do they collaborate in normal vertebrate brain development and function?** *BioEssays : news and reviews in molecular, cellular and developmental biology* 2007, **29:**1039-1047.
559. Cildir G, Akincilar SC, Tergaonkar V: **Chronic adipose tissue inflammation: all immune cells on the stage.** *Trends in molecular medicine* 2013, **19:**487-500.
560. Karin M: **Nuclear factor-kappaB in cancer development and progression.** *Nature* 2006, **441:**431-436.
561. Ghobrial IM, McCormick DJ, Kaufmann SH, Leontovich AA, Loegering DA, Dai NT, Krajnik KL, Stenson MJ, Melhem MF, Novak AJ, Ansell SM, Witzig TE: **Proteomic analysis of mantle-cell lymphoma by protein microarray.** *Blood* 2005, **105:**3722-3730.
562. Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, Alexe G, Lawrence M, O'Kelly M, Tamayo P, Weir BA, Gabriel S, Winckler W, Gupta S, Jakkula L, Feiler HS, Hodgson JG, James CD, Sarkaria JN, Brennan C, Kahn A, Spellman PT, Wilson RK, Speed TP, Gray JW, Meyerson M, et al: **Integrated genomic analysis identifies clinically**

relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer cell* 2010, **17:**98-110.

563. Dawson PJ, Wolman SR, Tait L, Heppner GH, Miller FR: **MCF10AT: a model for the evolution of cancer from proliferative breast disease.** *The American journal of pathology* 1996, **148:**313-319.
564. Choong LY, Lim S, Chong PK, Wong CY, Shah N, Lim YP: **Proteome-wide profiling of the MCF10AT breast cancer progression model.** *PloS one* 2010, **5:**e11030.
565. Choong LY, Lim SK, Chen Y, Loh MC, Toy W, Wong CY, Salto-Tellez M, Shah N, Lim YP: **Elevated NRD1 metalloprotease expression plays a role in breast cancer growth and proliferation.** *Genes, chromosomes & cancer* 2011, **50:**837-847.
566. Chan P, Mak M, inventors: **Blocking the migration or metastasis of cancer cells by affecting adhesion proteins the uses of new compounds thereof.** *US patent application 20,100,004,190* 2010.
567. Zucchi I, Mento E, Kuznetsov VA, Scotti M, Valsecchi V, Simionati B, Vicinanza E, Valle G, Pilotti S, Reinbold R, Vezzoni P, Albertini A, Dulbecco R: **Gene expression profiles of epithelial cells microscopically isolated from a breast-invasive ductal carcinoma and a nodal metastasis.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101:**18147-18152.
568. Liu Z, Li L, Yang Z, Luo W, Li X, Yang H, Yao K, Wu B, Fang W: **Increased expression of MMP9 is correlated with poor prognosis of nasopharyngeal carcinoma.** *BMC cancer* 2010, **10:**270.
569. Mitra RS, Goto M, Lee JS, Maldonado D, Taylor JM, Pan Q, Carey TE, Bradford CR, Prince ME, Cordell KG, Kirkwood KL, D'Silva NJ: **Rap1GAP promotes invasion via induction of matrix metalloproteinase 9 secretion, which is associated with poor survival in low N-stage squamous cell carcinoma.** *Cancer research* 2008, **68:**3959-3969.
570. Sakata K, Satoh M, Someya M, Asanuma H, Nagakura H, Oouchi A, Nakata K, Kogawa K, Koito K, Hareyama M, Himi T: **Expression of matrix metalloproteinase 9 is a prognostic factor in patients with non-Hodgkin lymphoma.** *Cancer* 2004, **100:**356-365.
571. Beliveau A, Mott JD, Lo A, Chen EI, Koller AA, Yaswen P, Muschler J, Bissell MJ: **Raf-induced MMP9 disrupts tissue architecture of human breast cells in three-dimensional culture and is necessary for tumor growth in vivo.** *Genes & development* 2010, **24:**2800-2811.
572. Nair RR, Avila H, Ma X, Wang Z, Lennartz M, Darnay BG, Boyd DD, Yan C: **A novel high-throughput screening system identifies a small molecule repressive for matrix metalloproteinase-9 expression.** *Molecular pharmacology* 2008, **73:**919-929.
573. Milde-Langosch K, Roder H, Andritzky B, Aslan B, Hemminger G, Brinkmann A, Bamberger CM, Loning T, Bamberger AM: **The role of the AP-1 transcription factors c-Fos, FosB, Fra-1 and Fra-2 in the invasion**

process of mammary carcinomas. *Breast cancer research and treatment* 2004, **86:**139-152.

574. Adamson R, Logan M, Kinnaird J, Langsley G, Hall R: **Loss of matrix metalloproteinase 9 activity in Theileria annulata-attenuated cells is at the transcriptional level and is associated with differentially expressed AP-1 species.** *Molecular and biochemical parasitology* 2000, **106:**51-61.
575. Chou YC, Sheu JR, Chung CL, Chen CY, Lin FL, Hsu MJ, Kuo YH, Hsiao G: **Nuclear-targeted inhibition of NF-kappaB on MMP-9 production by N-2-(4-bromophenyl) ethyl caffeamide in human monocytic cells.** *Chemico-biological interactions* 2010, **184:**403-412.
576. Ricca A, Biroccio A, Del Bufalo D, Mackay AR, Santoni A, Cippitelli M: **bcl-2 over-expression enhances NF-kappaB activity and induces mmp-9 transcription in human MCF7(ADR) breast-cancer cells.** *International journal of cancer Journal international du cancer* 2000, **86:**188-196.
577. Martinez N, Sanchez-Beato M, Carnero A, Moneo V, Tercero JC, Fernandez I, Navarrete M, Jimeno J, Piris MA: **Transcriptional signature of Ecteinascidin 743 (Yondelis, Trabectedin) in human sarcoma cells explanted from chemo-naive patients.** *Molecular cancer therapeutics* 2005, **4:**814-823.
578. Huang TT, Wuerzberger-Davis SM, Wu ZH, Miyamoto S: **Sequential modification of NEMO/IKKgamma by SUMO-1 and ubiquitin mediates NF-kappaB activation by genotoxic stress.** *Cell* 2003, **115:**565-576.
579. Mabb AM, Wuerzberger-Davis SM, Miyamoto S: **PIASy mediates NEMO sumoylation and NF-kappaB activation in response to genotoxic stress.** *Nature cell biology* 2006, **8:**986-993.
580. Lee MH, Mabb AM, Gill GB, Yeh ET, Miyamoto S: **NF-kappaB induction of the SUMO protease SENP2: A negative feedback loop to attenuate cell survival response to genotoxic stress.** *Molecular cell* 2011, **43:**180-191.
581. Ghosh A, Saginc G, Leow SC, Khattar E, Shin EM, Yan TD, Wong M, Zhang Z, Li G, Sung WK, Zhou J, Chng WJ, Li S, Liu E, Tergaonkar V: **Telomerase directly regulates NF-kappaB-dependent transcription.** *Nature cell biology* 2012, **14:**1270-1281.
582. Richter L, Bone JR, Kuroda MI: **RNA-dependent association of the Drosophila maleless protein with the male X chromosome.** *Genes to cells : devoted to molecular & cellular mechanisms* 1996, **1:**325-336.
583. Haas TL, Emmerich CH, Gerlach B, Schmukle AC, Cordier SM, Rieser E, Feltham R, Vince J, Warnken U, Wenger T, Koschny R, Komander D, Silke J, Walczak H: **Recruitment of the linear ubiquitin chain assembly complex stabilizes the TNF-R1 signaling complex and is required for TNF-mediated gene induction.** *Molecular cell* 2009, **36:**831-844.
584. Hoeller D, Hecker CM, Dikic I: **Ubiquitin and ubiquitin-like proteins in cancer pathogenesis.** *Nature reviews Cancer* 2006, **6:**776-788.
585. Chen ZJ: **Ubiquitin signalling in the NF-kappaB pathway.** *Nature cell biology* 2005, **7:**758-765.

586. Chen ZJ: **Ubiquitination in signaling to and activation of IKK.** *Immunological reviews* 2012, **246:**95-106.
587. Ring A, Dowsett M: **Mechanisms of tamoxifen resistance.** *Endocrine-related cancer* 2004, **11:**643-658.
588. Chabalier C, Lamare C, Racca C, Privat M, Valette A, Larminat F: **BRCA1 downregulation leads to premature inactivation of spindle checkpoint and confers paclitaxel resistance.** *Cell cycle* 2006, **5:**1001-1007.
589. Tanaka K, Kawaguchi H, Nakamura Y, Taguchi K, Nishiyama K, Ohno S: **Effect of HER2 status on risk of recurrence in women with small, node-negative breast tumours.** *The British journal of surgery* 2011, **98:**1561-1565.
590. Wu ZH, Miyamoto S: **Induction of a pro-apoptotic ATM-NF-kappaB pathway and its repression by ATR in response to replication stress.** *The EMBO journal* 2008, **27:**1963-1973.
591. Low KC, Tergaonkar V: **Telomerase: central regulator of all of the hallmarks of cancer.** *Trends in biochemical sciences* 2013, **38:**426-434.
592. Qian P, Zuo Z, Wu Z, Meng X, Li G, Wu Z, Zhang W, Tan S, Pandey V, Yao Y, Wang P, Zhao L, Wang J, Wu Q, Song E, Lobie PE, Yin Z, Zhu T: **Pivotal role of reduced let-7g expression in breast cancer invasion and metastasis.** *Cancer research* 2011, **71:**6463-6474.
593. Pandey V, Jung Y, Kang J, Steiner M, Qian PX, Banerjee A, Mitchell MD, Wu ZS, Zhu T, Liu DX, Lobie PE: **Artemin Reduces Sensitivity to Doxorubicin and Paclitaxel in Endometrial Carcinoma Cells through Specific Regulation of CD24.** *Translational oncology* 2010, **3:**218-229.
594. Miller LD, Coffman LG, Chou JW, Black MA, Bergh J, D'Agostino R, Jr., Torti SV, Torti FM: **An iron regulatory gene signature predicts outcome in breast cancer.** *Cancer research* 2011, **71:**6728-6737.
595. Edgar R, Domrachev M, Lash AE: **Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.** *Nucleic acids research* 2002, **30:**207-210.
596. Pawitan Y, Bjohle J, Amler L, Borg AL, Egyhazi S, Hall P, Han X, Holmberg L, Huang F, Klaar S, Liu ET, Miller L, Nordgren H, Ploner A, Sandelin K, Shaw PM, Smeds J, Skoog L, Wedren S, Bergh J: **Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts.** *Breast cancer research : BCR* 2005, **7:**R953-964.
597. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, Ellis P, Harris A, Bergh J, Foekens JA, Klijn JG, Larsimont D, Buyse M, Bontempi G, Delorenzi M, Piccart MJ, Sotiriou C: **Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade.** *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007, **25:**1239-1246.
598. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, Gillet C, Ellis P, Ryder K, Reid JF, Daidone MG, Pierotti MA, Berns EM, Jansen MP, Foekens JA, Delorenzi M, Bontempi G, Piccart MJ, Sotiriou C: **Predicting**

prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC genomics* 2008, **9:**239.

599. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP: **Summaries of Affymetrix GeneChip probe level data.** *Nucleic acids research* 2003, **31:**e15.
600. Johnson WE, Li C, Rabinovic A: **Adjusting batch effects in microarray expression data using empirical Bayes methods.** *Biostatistics* 2007, **8:**118-127.